THE ANOMERIC EFFECT IN ENZYME-CATALYZED PHOSPHORYL TRANSFER

By

Jean C. Summerton

A Dissertation

Presented to the Department of Biochemistry & Molecular Biology

and the Oregon Health & Science University

School of Medicine

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

April 2014

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

_____

This is to certify that the PhD dissertation of

Jean Camille Summerton

has been approved

_____
Michael Chapman

_____
William Skach

_____
Eric Gouaux

_____
Francis Valiyaveetil

_____
Larry David

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figures in Introduction

Figures in Chapter 1

Figures in Chapter 2

Figures in Conclusions

LIST OF NON-STANDARD ABBREVIATIONS USED

QM	Quantum mechanics
MM	Molecular mechanics
AK	Arginine Kinase
NTP	Nucleotide triphosphate
PC	Phosphocreatine
n(O)	Lone pair orbital on an oxygen atom
σ*	Anti-bonding orbital
DMSO	Dimethyl sulfoxide
PCM	Polarizable continuum model
NBO	Natural bond orbitals
E(2)	2^{nd} order perturbation energy – a measure of hyperconjugation.
$F_{i,j}$	Fock matrix element – a measure of orbital overlap.
E_j-E_i	Orbital energy gap
n(O) → σ*	Donation of electron density from a lone pair orbital into an anti-bonding orbital.
GAP	GTPase activating protein
μ_σ	Standard error of the mean
p-loop	Phosphate-binding loop
FTIR	Fourier transform infrared spectroscopy

This work is dedicated to my family. Without their support this would not have been possible. Thank you, James, Daniel and Cirelle Summerton and Marisa Schwartzman. Especially I want to thank my parents, Jim and Patricia Summerton whose love and encouragement have made me the person I am today.

ACKNOWLEDGMENTS

First, I would like to thank my mentor, Dr. Michael Chapman. Michael's mentoring style constitutes the perfect balance of letting you learn from your own mistakes, while at the same time, always being there for questions and support. Thank you especially for your seemingly unending patience. I couldn't have joined a better lab.

I would also like to thank the current and past members of the Chapman lab, and the Evanseck lab, with particular appreciation given to Dr. Sai Pakkala for his commiseration and friendship.

Thank you to my committee, Drs. Bill Skach, Eric Gouaux, Francis Valiyaveetil and Larry David. Especially thank you Eric and Bill for not minding when I drop by, unannounced, for support and advice. I very much appreciate everything you have done.

Thank you to the Biochemistry and Molecular Biology department and especially the Biochemistry graduate students and postdocs. Our close-knit department made this whole PhD. thing a breeze… My heartfelt thanks in particular go out to Dr. Samantha Gardner, Nancy Meyer, Danielle Williamson, Johannes Elferich, Jessica Martin, Andrew Hadd, Gregory Martin and Jeni Wroblewski.

Thank you Lizzy Scheibel, for listening to ALL of my practice presentations. I promise you will never have to hear about the anomeric effect again. Finally thank you Jeremy Bushman, for being there for the stress and joys.

ABSTRACT

Phosphoryl transfer is ubiquitous in biology. Defects in the enzymes that catalyze such transfer upset normal cellular energy metabolism, motility and the generation of transmembrane potentials. In “high energy” phosphoanhydride bonds, such as those in ATP and GTP, computational evidence has emerged showing a dependence of phosphoanhydride bond energy and reactivity on the anomeric effect. The hypothesis guiding this dissertation research is that enzymes have evolved to take advantage of an intrinsic anomeric effect in nucleotide ligands to drive catalysis.

The first section addresses how enzymes might modulate the anomeric effect for catalytic gain. The potential impact of hydrogen bonding from neighboring molecules is calculated using high level quantum mechanical analysis. It is found that hydrogen bonds between water and the terminal phosphoryl group of an ATP analog induce a reduction in the anomeric effect and a shortening/strengthening of the scissile phosphoanhydride bond. Although the anomeric effect requires strict orbital alignments, geometrical requirements for hydrogen bonds are minimal. Thus a larger array of interactions than previously thought has the potential to modulate the strength of the scissile bond.

The second part of the dissertation addresses whether enzymes have evolved the potential to modulate bond strength through the anomeric effect. A top-down analysis of protein Data Bank structures was performed to examine active site interactions that might be prevalent in phosphoryl transfer enzymes, the hypothesis being that these enzymes might have interactions that would optimize hyperconjugation. The database survey highlighted several key interaction patterns. Particularly abundant in phosphoryl transfer

enzymes, were hydrogen bonds with the bridging oxygen of the scissile phosphoanhydride bond. Computation shows these interactions elongate the scissile bond by enhancing the anomeric effect. Experimental mutagenesis against highlighted interactions in arginine kinase resulted in a 100-fold decrease in catalytic rate, providing both evidence for a relationship between the anomeric effect and catalytic rate and support for the theory that the anomeric effect is a factor in AK catalysis.

There is increasing appreciation that enzymes can employ multiple mechanisms of rate enhancement. This dissertation research uncovered a hitherto overlooked mechanism to accelerate phosphoryl transfer that has not yet proven to be the dominant means of rate enhancement in any enzyme, but appears to be an important component in many enzymes, establishing a common denominator between families of enzymes that were previously thought to operate with quite distinct mechanisms of action.

INTRODUCTION

Phosphoryl transfer enzyme catalysis in Arginine kinase – Catalysis of reactions involving high energy phosphoanhydride bonds, such as those in ATP, is central to most cellular processes. Defects in catalysis of phosphoanhydride bonds can lead to life-threatening conditions.[1]

Creatine kinase, found in metazoa, buffers ATP levels in myocytes, neurons and other cells. Creatine kinase carries out a bimolecular reaction where a phosphate from ATP is reversibly transferred to creatine to make ADP and phosphocreatine. Although this is an unfavorable reaction (equilibrium constant of 0.012 at pH 7.4 with 1 mM free Mg^{2+}), the high MgATP/MgADP ratio in resting muscle allows build up of a creatine phosphate/creatine ratio of ~2, so that MgATP, depleted by muscle contraction, can be rapidly regenerated by creatine kinase.[2] Due to their importance in cell function, phosphogen kinases such as creatine kinase, have been classic enzymological models of multisubstrate reactions.[3]

Arginine kinase is an invertebrate homolog of creatine kinase that catalyzes transfer of the terminal phosphate of ATP to arginine, creating phosphoarginine and ADP. Arginine kinase is active as a monomer, making it particularly amenable to biophysical and biochemical characterization relative to creatine kinase (dimer or octamer). Arginine kinase shares the same catalytic mechanism and conserved active site features as creatine kinase [4,5], making it an optimal model for study of bimolecular transferase catalysis.

Early mechanistic proposals focused on general base catalysis, in which the substrate arginine would be deprotonated prior to its attack on the ATP. A 1.2 Å

transition state analog complex of arginine kinase showed only two glutamate residues sufficiently close to the arginine substrate and not the histidine implicated as the catalytic base by pH-dependent studies .[6] The pK of Glu suggested that glutamates would be most effective in proton abstraction after formation of the new N—P bond, rather than before, as had been presumed for the suggested histidine. However, mutagenesis of the glutamates showed that at most they played an accessory role in catalysis.[7] It was concluded that the Glu residues might also be important for proper alignment of substrates. Further mutagenesis work on active site residues in arginine kinase pointed to a Cys thiolate as potentially important for electrostatic stabilization of the transition state.[8]

A search commenced to identify additional factors to more fully account for overall catalytic enhancement. Initial efforts were focused on the chemistry of phosphates, to gain insights into how enzymes may manipulate nucleotide chemistry to achieve catalysis.

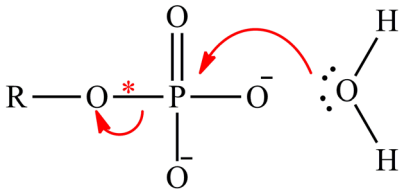


Figure 1. Phosphoryl transfer. Arrows indicate transfer of electrons. The scissile bond is shown with a red asterisks

Chemistry of phosphoryl transfer in solution – Phosphates undergo nucelophilic displacement reactions where a nucleophile; a water in hydrolysis reactions and another substrate in phosphoryl transfer, attacks at the electropositive phosphate atom. Electrons from the scissile O—P bond are transferred to the (previously) bridging oxygen upon

bond cleavage and the phosphate is displaced from the R—O leaving group (Figure 1).

Nucelophilic displacement reactions generally require a good leaving group. Synthetic organic chemists commonly use chlorides, bromides, tosylates or triflates however, they almost never use the weaker phosphate leaving groups that are common in nature.[9] Before considering the details of catalysis, we might ask why phosphates are so ubiquitous in biochemistry, particularly when far better leaving groups for nucleophilic substitution are available?[9] Westheimer points to several chemical characteristics;[9] 1) The low pka of phosphates keeps them in ionized form at cellular pH, thus preventing significant loss through the cell membrane; 2) The unique chemistry of phosphates makes them stable in solution while at the same time being susceptible to incredible rate enhancement by enzymes. Phosphoryl transfer in solution has a general rate constant of approximately 10^9 to 10^{12} fold less than enzyme-catalyzed reactions,[9] allowing the body to safely transport phosphates and activate them when necessary. This property, often referred to a kinetic stability coupled with thermodynamic instability, [10] is critical for integrity of our genetic material, distribution of metabolites throughout the cell, and the cell's use of phosphate groups as key signaling factors. It is thought that the negative charges on phosphates partially prevent nucelophilic attack by waters in solution.[11]

While phosphates must remain intact in a solution environment so too must they retain attributes that may be manipulated by enzymes to enhance reaction rates. What properties of phosphates make them ideal for enzymatic rate enhancement? The answer to this question may lay in the chemistry of the "high energy" phosphoanhydride bonds.

The reactivity of high energy bonds – High-energy bonds are those that release substantial free energy upon hydrolysis ($\Delta G°' < -25$ kJ/mol). As detailed below, three

factors are usually cited in the high energy character of these bonds: differences in solvation, competing resonance and electrostatic repulsion.[12] By differential solvation, it is meant that the hydrolysis products can be solvated more favorably than the reactant phosphoanahydride. In competing resonance theory, the hydrolyzed products have fewer bonding orbitals competing for lone pair electrons than does the phosphoanhydride reactant, where the lone pairs on the bridging oxygen are drawn to both phosphate groups (Figure 2). Thus, the phosphoanhydride is destabilized relative to the hydrolyzed products. Electrostatic repulsion occurs between adjacent negatively charged phosphates. It is also thought the profusion of electron density associated with the phosphoryl groups allows for large differentials in electrostatic stabilization, making it possible for enzymes to fine tune reaction paths with only slight changes to active site environment.[13]

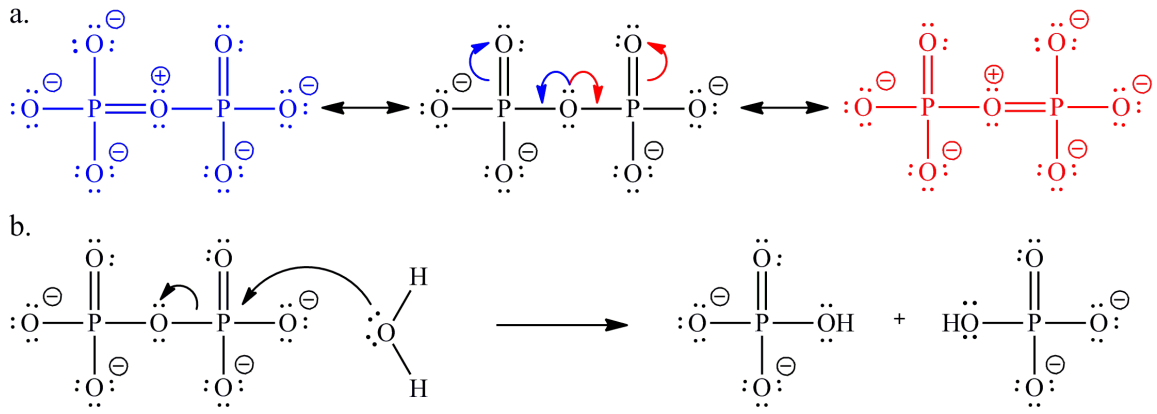


Figure 2. Competing resonance a) Competition between multiple potential resonance structures is destabilizing for the molecule. b) The more stable products of the hydrolysis reaction allow for greater total resonance.

Although these explanations appear to provide a qualitative rationale for the reactivity of ATP, they fail to fully explain the reactivity of other high energy compounds. For example, in phosphoguanidino compounds such as phosphoarginine and

phosphocreatine, the hydrolysis products are oppositely charged (Figure 3), i.e. attractive, so charge repulsion is not a factor even though the enthalpies of hydrolysis are commensurate with those of nucleotides. Furthermore, competing resonance has been shown to be only a minor contributor to the energy.[14] Such shortfalls in our understanding prompted a search for missing components to explain the high energy character of these bonds. It was found that stereoelectronics, as discussed below, can account for high energy bond reactivity in a broad range of compounds.[14,15]

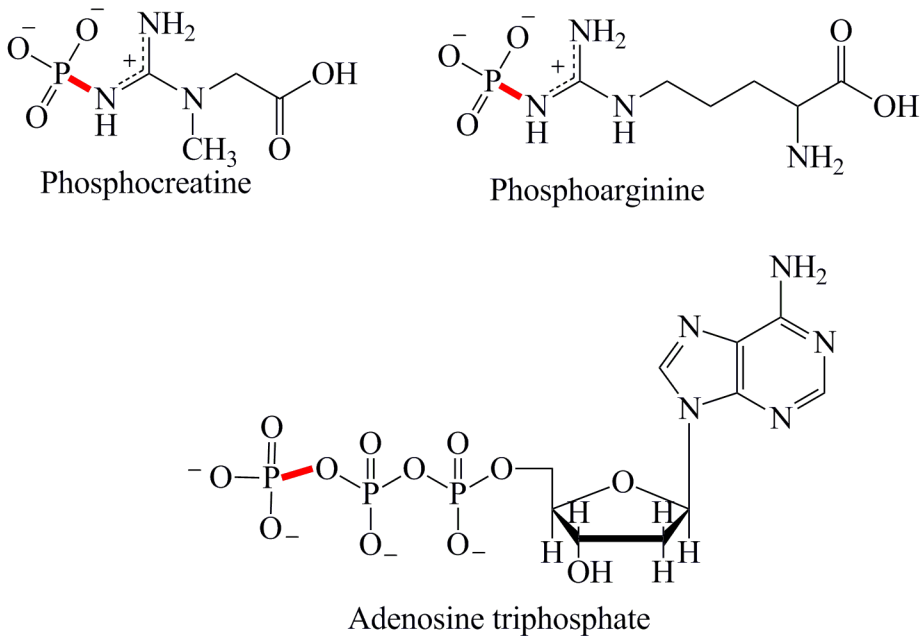


Figure 3. Structures of phosphocreatine, phosphoarginine and ATP. The high energy bond is shown in red.

The anomeric effect – Stereoelectronic effects are defined as the effects of electronic structure on the reactivity and physical properties of a molecule. Interactions between atomic and molecular orbitals can drive preferences in conformation, resonance structure and transition states, leading to distinct reaction selectivities.[16]

The anomeric effect is the transfer of electron density from a lone pair or bonding orbital into an anti-bonding orbital of another bond. The anomeric effect is a type of stereoelectronic effect where the conformational energy depends upon the alignment of orbitals for a favorable interaction.[16] It was first recognized by J.T. Edward in 1955 and R.U. Lemieux in 1958 as a determinant of the configuration of hydroxyl groups in sugars.[17,18] For example, β-D-glucopyranose prefers an axial orientation rather than the less hindered equatorial orientation that would be otherwise expected (Figure 4). In the favored configuration, the oxygen lone pair orbital (n) is anti-periplanar relative to the hydroxyl, and therefore *cis* and overlapping with the hydroxyl antibonding orbital (σ*), establishing a favorable electronic interaction.[19] In addition to favoring this conformer, the anomeric interaction leads to a weakening / lengthening of the C—OH bond.

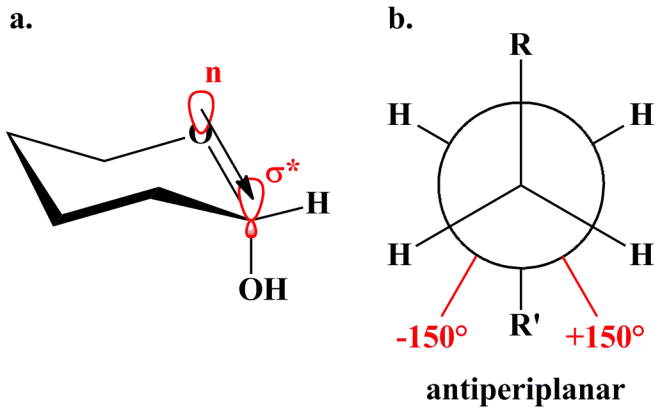


Figure 4. a) The oxygen lone pair orbital is anti-periplanar relative to the hydroxyl, and therefore adjacent to and overlapping its antibonding orbital (σ), establishing a favorable electronic interaction. b) Newman projection where stereochemical arrangements corresponding to torsion angles of R and R' between +/- 150° to 180° are called antiperiplanar. Figures adapted from IUPAC Compendium of Chemical Terminology.*

In more recent times "anomeric effect" has been used in a more general context to refer to configuration dependent interactions between lone pair or bonding electrons with an aligned anti-bonding orbital (σ*) (Figure 5).

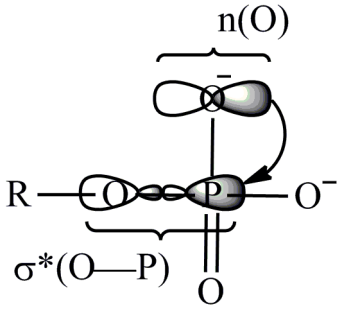


Figure 5. Rendering of the anomeric effect in a phosphate where electron density is transferred from the oxygen lone pair (n(O)) into the antibonding orbital of the O—P bond (σ(O—P)).*

The potential for a large effect in phosphoryl transfer catalysis was realized as a possibility only after Gorenstein's seminal work on a range of high energy compounds.[20-22] An important role for the anomeric effect was further established with this lab's studies of high-energy phospho-metabolites.[15] Here, high level quantum mechanical calculations on a series of 10 different phosphate-containing compounds showed that the energy and length of reactive O—P bonds are strongly correlated to the magnitude of anomeric interactions between phosphate oxygen lone pairs and an anti-bonding orbital (σ*) of O—P bonds. As the magnitude of anomeric interaction decreases, the bond is strengthened/shortened. The degree of anomeric interaction was also found to correlate strongly with experimental energies of hydrolysis supporting the proposed importance of the anomeric effect in reactions.

Little is known about the impact of environment on the anomeric effect. – If enzymes are to successfully modulate the anomeric effect for catalytic gain, the anomeric

effect itself must be mutable through interactions between a reactant and its environment. There is some evidence to suggest that this is the case. Precise IR experiments have demonstrated an increase in scissile O—P bond lengths in going from Water to DMSO.[23] This finding was supported by small molecule quantum studies.[15] It was found that the more polar solvent reduces the anomeric effect, explaining both the calculated and experimentally observed changes in scissile bond length. Water was thought to sequester lone pair electron density making it unavailable for anomeric interactions.

As a water environment decreases scissile bond length and the anomeric effect, it follows that enzymes might mitigate the effects of solvent and might even enhance the anomeric effect above what could be expected in the gas phase. How enzymes enhance the anomeric effect over solution state level is not yet known.

Current understanding of phosphoryl transfer enzyme mechanisms – Much has been learned about enzyme-catalyzed phosphoryl transfer through the detailed study of individual enzymes. Frequently cited mechanisms include; a) positioning of substrates for phosphate transfer [24-29], b) base-assisted activation of the nucleophile [30-34], c) activation of the electrophile [28,35-37] and finally d) electrostatic stabilization [27,33,37-40] of the transferring phosphate group (Figure 6).

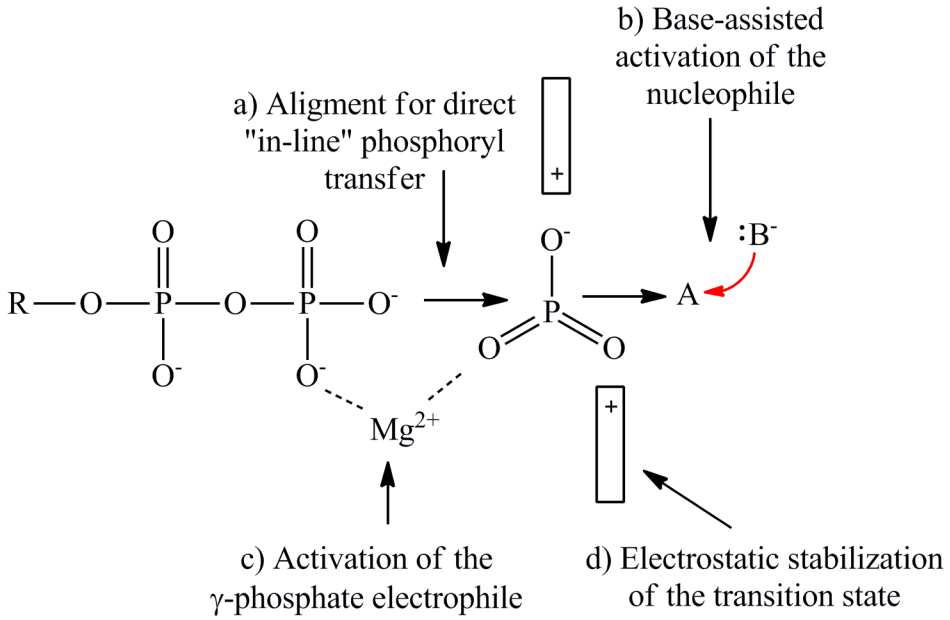


Figure 6. Frequently cited mechanisms of enzyme-catalyzed phosphoryl transfer. "A" represents the phosphoryl group acceptor and bars represent positively charged amino acid side chains. a) Pre-alignment of the phosphoryl group donor and acceptor, brings the substrates into a near-attack configuration, reducing the free energy of attaining the transition state relative to free solution [41]. b) A general base deprotonates the attacking nucleophile, activating it for attack on the γ-phosphate [12]. Metal cations may also polarize the nucleophile. The γ-phosphate is also activated through withdrawal of electron density by c) metal ions or d) positively charged amino acids. Such interactions are also thought to help stabilize negative charge on the transferring phosphate group in the transition state.

Debate has continued over the role of base-assisted catalysis in various systems [26,28,37,42] and there have been additional mechanisms proposed for specific enzyme types. Proton relays have been proposed, where the attacking water is deprotonated by another water distant from the nucleotide (Figure 7). [12,28,43,44] Substrate-assisted catalysis [42,45] in

conjunction with electrostatic stabilization of the transition state has been championed by the Warshel group in particular [46-48]. Here the γ-phosphate itself acts as the general base. Finally neutralization of charge repulsion between reactants [28,39,43,49,50], strain of the β and γ phosphate groups to destabilize the nucleotide reactant [33,49,50] are also cited.

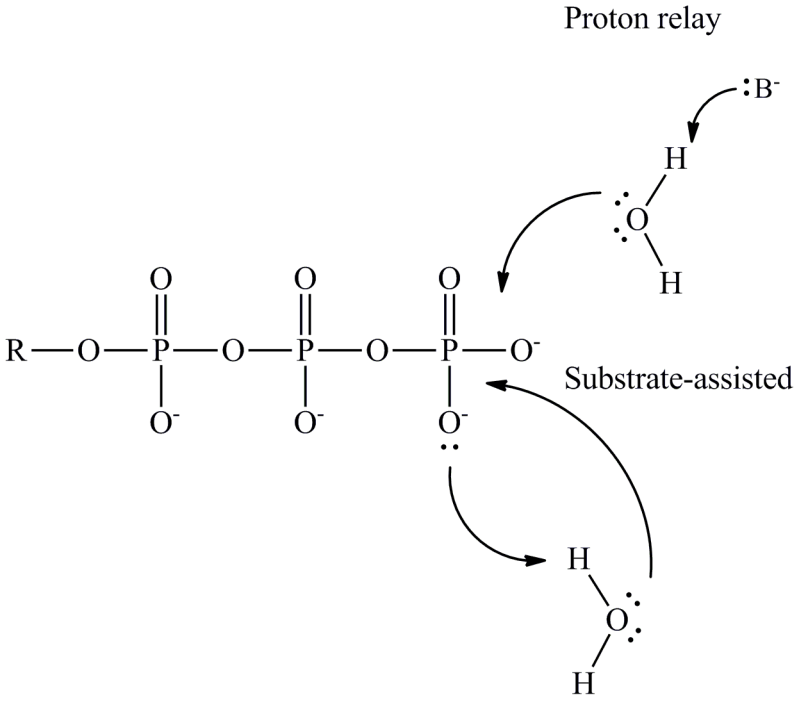


Figure 7. Debate over the plausibility of base-assisted catalysis in several enzyme mechanisms sparked new proposals. In proton relay, a base (B) deprotonates the nucleophile water. The γ-phosphate deprotonates the water in substrate-assisted catalysis. Arrows are meant to indicate the general movement of electrons but do not represent the formal chemical reactions.

Phosphoryl transfer enzymes are a broad class within which several mechanisms may have evolved. However, one is struck by the diversity in mechanistic proposals and the lack of consensus on elements of mechanism that might be shared between enzyme families.

Phosphates such as ATP have unique chemistries; particularly striking are the extensive electron clouds surrounding the phosphoryl oxygens. In the proposed phosphoryl transfer mechanisms, little advantage is taken of specific phosphate chemical characteristics (one exception to this is electrostatic stabilization of the transition state[46]). It begs the question as to whether contributions to mechanism may have been overlooked because they did not have well established precedents in other enzymes. It is known that the extensive reserve of electron density in ATP leads to an exceptional anomeric effect. Here, it is proposed that the importance of ATP in biology stems from this property and that enzymes have evolved to take advantage of the anomeric effect for catalytic gain.

Evidence for the anomeric effect in Arginine kinase – Does the anomeric effect play a role in catalysis in arginine kinase (AK)? Unpublished work by a former graduate student in the Chapman lab investigated this question using computation.[51] A simulation was designed using just the active site residues and (replacing the ADP and NO_3 of the analog with substrate ATP). Interacting residues were removed (computationally) one-by-one and then the energetics and geometry of the system recalculated. A strong correlation was found between the anomeric effect and the computed O—P bond lengths. These preliminary computations suggested that the enzyme could be modulating reaction chemistry through the anomeric effect.

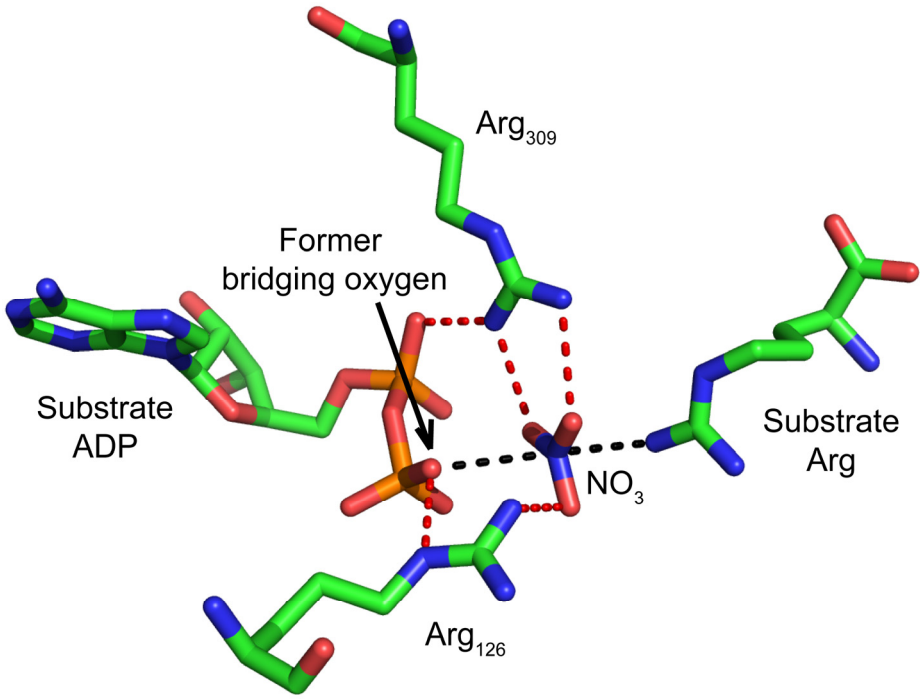


Figure 8. Select residues in Arginine Kinase active site. Black dotted lines indicate the path of in-line phosphate transfer and red lines denote hydrogen bonds.

It was found that arginine residues R309 and R229 cause a shorter, more stable scissile O—P bond, yet other arginine residues cause a longer (Arg 126), less stable O—P bond. The location of R309 and R126 in relation to ligands in the TSA complex of arginine kinase (PDB ID 1M15[52]) are shown in Figure 8. In the TSA structure, the NO_3 is an analog for the γ-phosphate group in its presumptive transition state position. Arg_{126} makes one contact to the formerly bridging oxygen of the nucleotide and one contact to the γ-phosphate analog. Arg_{309} makes two contacts to the γ-phosphate analog. It was suggested, but not demonstrated, that the contact between Arg_{126} and the bridging oxygen of the scissile bond might increase the anomeric effect by polarizing the O—P bond. Polarization of the O—P bond is known to enhance the anomeric effect.[15] As the anomeric effect relies on proper alignment of orbitals, an alternative explanation is that

R_{126} promotes a ligand conformation that maximizes the anomeric effect. These preliminary computational studies provided a) evidence that enzymes could manipulate the anomeric effect, perhaps for catalytic gain, and b) clues as to how active site residues might accomplish this through specific hydrogen bonding interactions.

Summary of the dissertation research – Previous modeling studies in the Chapman lab have provided strong evidence for the relationship between the anomeric effect and the chemistry of high energy O—P and N—P bonds[14,15] in individual molecules. The work presented in this dissertation shifts the focus to enzymes, and looks at how active sites manipulate the anomeric effect for catalytic gain. Three Questions pertinent to the role of the active site in modulating the anomeric effect are discussed below.

1. *What types of Molecular Interactions enhance the anomeric effect?* In Chapter one, high level quantum mechanics are used to determine how the anomeric effect in ATP might be reduced through interaction with other molecules. A continuation of this work is given in Chapter two, where interactions are investigated that enhance the anomeric effect.
2. *Do a broad range of enzymes exploit the anomeric effect?* Using database analysis of Protein Data Bank crystal structures, the prevalence was investigated of enzyme active site configurations that might promote anomeric destabilization.
3. *To what degree does the anomeric effect enhance rate?* Much of this dissertation research uses changes in bond length as an indicator of the potential impact on the reaction. Other factors contribute to rate enhancement. Thus, the final chapter of the research measures experimentally the impact of perturbing one of the

previously discovered interactions through steady state kinetics following mutation in one model enzyme system. The results serve as at least a preliminary experimental corroboration that the computational investigation of the anomeric effect is relevant to the real workings of enzymes.

The work presented in this dissertation investigates an important element of rate-enhancement that has been largely unexplored. Results are shown to be relevant to a substantial fraction of a broad class of enzymes that are ubiquitous in biology.

1. HYPERCONJUGATION-MEDIATED SOLVENT EFFECTS IN PHOSPHOANHYDRIDE BONDS

Reproduced with permission from Summerton J.C., Evanseck J.D., Chapman M.S. (2012) J Phys Chem A 116: 10209-10217.

ABSTRACT

Density functional theory and Natural Bond Orbital analysis are used to explore the impact of solvent on hyperconjugation in methyl triphosphate, a model for "energy rich" phosphoanhydride bonds, such as found in ATP. As expected, dihedral rotation of a hydroxyl group vicinal to the phosphoanhydride bond reveals that the conformational dependence of the anomeric effect involves modulation of the orbital overlap between the donor and acceptor orbitals. However, a conformational independence was observed in the rotation of a solvent hydrogen bond. As one lone pair orbital rotates away from an optimal *anti*-periplanar orientation, the overall magnitude of the anomeric effect is compensated approximately by the other lone pair as it becomes more *anti*-periplanar. Furthermore, solvent modulation of the anomeric effect is not restricted to the *anti*-periplanar lone pair; hydrogen bonds involving *gauche* lone pairs also affect the anomeric interaction and the strength of the phosphoanhydride bond. Both *gauche* and *anti* solvent hydrogen bonds lengthen non-bridging O—P bonds, increasing the distance between donor and acceptor orbitals, and decreasing orbital overlap which leads to a reduction of the anomeric effect. Solvent effects are additive with greater reduction in the anomeric effect upon increasing water coordination. By controlling the coordination environment of substrates in an active site, kinases, phosphatases and other enzymes important in

metabolism and signaling, may have the potential to modulate the stability of individual phosphoanhydride bonds through stereoelectronic effects.

INTRODUCTION

ATP is vital to energy exchange in biological systems. The free energy liberated in the hydrolysis of ATP derives from "energy-rich" phosphoanhydride bonds that are essential for maintaining life.[53] While crystallography,[52,54,55] biochemical kinetics,[56-59] thermodynamic studies [60,61] and computation[62-64] have all furthered the understanding of phosphate ester hydrolysis, the chemical basis of "high-energy" or "energy-rich" O—P bonds continues to be debated.[65-67] Our incomplete description of phosphanhydride chemistry in molecules, such as ATP and phosphocreatine (PC),[15] limits our understanding of the enzymology of phosphoryl transfer in a wide variety of kinases and phosphatases, like creatine kinase, myosin and other ATPases.

The lability of the O—P bond in molecules like ATP is traditionally attributed to three factors:[12,53] relative to the hydrolyzed products, the ATP reactant has less resonance stabilization,[68] greater electrostatic repulsion[69] and less stabilization due to solvation.[70] However, the nature of the O—P phosphoanhydride bonds in ATP has been re-evaluated recently and stereoelectronic effects were found to be significant.[15,71] An anomeric effect appears to be more important than other delocalization interactions in the length and implied strength of O—P bonds.[15]

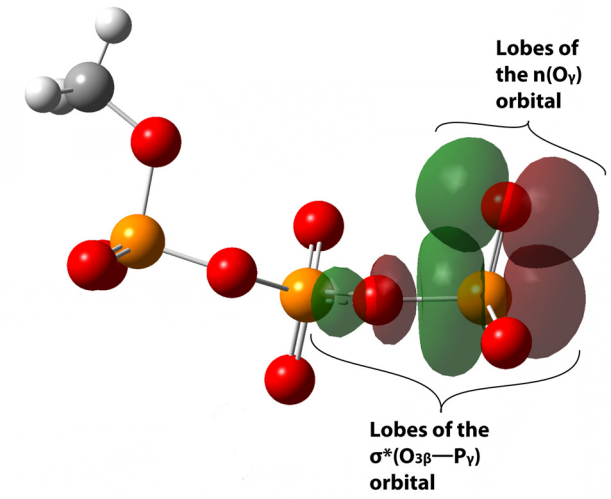


Figure 9. Anomeric effect written as $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$ is defined as the transfer of electron density from a lone pair orbital on a γ-oxygen ($n(O_\gamma)$) to the anti-bonding orbital of the $O_{3\beta}—P_\gamma$ bond ($\sigma^*(O_{3\beta}—P_\gamma)$). The orbitals are colored green and purple with arbitrary phase.

The O—P bond is destabilized or lengthened by a specific type of hyperconjugation that is popularly referred to as the anomeric effect.[15] The anomeric effect is transfer of electron density from a lone pair or bonding orbital (donor orbital) into an anti-bonding orbital (acceptor orbital) of another bond. Anomeric interactions have been found to be maximal when the donor orbital is *anti*-periplanar to the vicinal bond, so that the donor orbital is in close proximity to the antibonding orbital of the neighboring bond for efficient electron density transfer between the two.[22,72,73] One such interaction in a triphosphate is shown in Figure 9. The lone pair (donor) on a gamma oxygen interacts with the antibonding orbital (acceptor) of the sigma bond between the gamma phosphorous and the adjacent beta oxygen, and is denoted as the $n(O_\gamma) \rightarrow \sigma^*(O_\beta—P_\gamma)$ hyperconjugation. The term "anomeric effect" has been used to describe stereo-preferences of sugars[18] and related molecules.[74] Recent work has argued that

stereo-preferences are actually a result of destabilizing electrostatic interactions.[75] In this work the role of hyperconjugation in phosphodiester bond length is explored. The current terminology of describing some of the interaction as an "anomeric effect" is used following prior suggestions that the hyperconjugation was similar to that thought earlier to be a determinant of sugar conformation. In following this historical terminology, no implications are intended about the actual determinants of sugar conformation.

Strain of a substrate towards the ideal *anti*-periplanar dihedral angle that maximizes the anomeric effect and destabilizes the bond has been proposed as a potential mechanism of enzyme catalysis, termed torsional activation. Gorenstein *et al.* explored the effects of anomeric bond destabilization in the formation of peptide N—C bonds in the gas phase reaction between ammonia and formic acid, chosen to model serine proteases.[20] The study revealed that a lone pair on the hydroxyl oxygen, which is *anti*-periplanar to the attacking nitrogen, stabilized the transition state by 3.9 kcal/mol. They also examined O—P bond reactivity in the gas phase: in the dimethyl phosphate monoanion, there is a weakening by ~1 kcal/mol when oxygen atom lone pairs are *anti*-periplanar toward the bond *versus gauche*.[21]

A majority of quantum calculations have been performed *in vacuo*, inherently omitting environmental effects that could conceivably modulate stereoelectronic effects.[76-79] The effect of solvent on conformational energetics was demonstrated in quantum studies of the phosphodiester links of DNA.[80] In NMR-restrained molecular dynamics studies of DNA and the DNA-binding domain of the transcription activator protein, PhoB, it was found that the solution structure provided better local geometry and convergence than structures determined in vacuum.[81] Recent studies with implicit

solvent found strong correlations between the anomeric effect and O—P and N—P bond lengths in analogs of nucleotides and phosphocreatine.[14,15] A reduction in anomeric effect was observed when lone pair donor ability was decreased (with solvent) or the electron density acceptor ability of the *anti*-bonding acceptor was decreased.

The influence of environmental factors or conformation upon the anomeric effect remains an open question germane to potential enzyme mechanisms.[14,15,20,82] Yang *et al*. recently used QM/MM calculations to analyze ATP in the open (post-rigor) and closed (pre-powerstroke) active sites of the myosin motor domain.[83] Despite a substantial difference in the scissile O—P bond length (0.05 Å), Arg_{238}, the residue with the largest effect on the anomeric interaction, had only modest impact on the ground state energy. Rudbeck *et al*. used a similar approach to look at the environmental impact of the Ca^{2+}-ATPase on its phosphoenzyme intermediate and found that in the inactive protein model, the NH_3 group of Lys_{684} induces shortening of the P−O_β bond due to changes in hyperconjugation, while in the inactive model, the P—O_β bond is elongated.[84] Both of these studies looked at destabilization of the ground state, however, other potential catalytic effects, such as transition state stabilization, remain to be investigated. Although hyperconjugation lengthens the scissile bond, paradoxically, it also stabilizes the molecule as a whole. Potentially then, hyperconjugation could lower the energy of the transition state while also destabilizing the scissile bond.

The rationale for analyzing hydrogen-bonding interactions stems from our prior finding that the anomeric effect depends on protonation state of the γ-phosphate in methyl triphosphate.[15] Involvement of the gamma phosphate oxygen lone pair in a hydrogen bond reduces the electron density available for transfer in an anomeric

interaction.[15] The potential importance of hydrogen bonding interactions was also highlighted by Iche-Tarret *et al*. who found that an intramolecular hydrogen bond in protein tyrosine phosphatase (PTPase) substrate analogs destabilized the scissile O—P bond.[85] In an enzyme active site, both protein and solvent atoms could form relevant hydrogen bonds, but interactions with water molecules are of particular interest. A number of kinases undergo large conformational changes, to exclude water from the active site, perhaps to prevent premature hydrolysis of the ATP substrate.[86] By contrast, the active sites of some enzymes contain water molecules in conserved locations [87] that can have multiple roles in catalysis.[88-90] Thus, we chose the interactions of water with methyl triphosphate both as a general model for the interactions of polar groups with phosphoanhydride bonds, and also to provide insights into the diverse roles of water upon hyperconjugation in phosphoryl transfer active sites.

Despite its crucial role in physiological processes, the factors governing O—P phosphoryl bond weakening and possible connection to ”high energy” character of phosphates remains controversial.[23,91-94] The current work explores how molecular conformation and solvent interactions impact ground state O—P bond weakening through stereoelectronic effects. We show that water can have a strong impact on phosphoanhydride bond length. Changes to the solvation upon substrate-binding, and, by inference, hydrogen bonding with other polar active site groups, have the capacity to alter significantly the energetics of catalysis, and the specificity of which phosphoanhydride bond would be destabilized in a particular enzyme-catalyzed reaction.

METHODS

Electronic structure calculations were carried out with the Gaussian 09 suite.[95] The B3LYP/6-311++G(d,p) level of theory was used for structure optimization.[82,96-98] Prior studies[15] showed that this level of theory captured the same hyperconjugative effects as second-order Moller-Plesset[99] computations. In addition to the water molecules explicitly modeled, the effects of bulk solvent were modeled implicitly using a polarizable continuum model (PCM).[100]

Combining explicit waters with continuum approach such as PCM can lead to inaccuracies in handling the polarization of the first solvation shell and entropic terms in free energy calculations.[101] However for potential energy calculations used to assess enthalpies of explicit interactions, a mixed model has been shown to be an appropriate treatment.[102] Thus, in this work, nearest-neighbor explicit water molecules are embedded in a continuum approximation.[103]

Natural bond orbital (NBO) analysis was performed using the GenNBO5 program[104] with output files from the Gaussian 09 program. HF/cc-PVTZ model chemistry was used for all NBO calculations upon B3LYP/6-311++G(d,p) geometry optimized structures.[105] NBO transforms the non-orthogonal atomic orbitals from the HF wavefunction into natural bond orbitals (NBO)[106] allowing for a Lewis-like description of the total N-electron density[107]. Delocalizing interactions can be treated through second-order perturbation theory. The second order perturbation (E(2)) energy values provide an estimate of the magnitude of delocalizing interactions and are calculated by[108]

$$E(2) = q_i \cdot \frac{F(i,j)^2}{E(j)-E(i)} \quad (1)$$

where q_i is the donor occupancy. The difference in energy of the donor and acceptor orbitals is E(j)-E(i) and the orbital interaction energy is F(i,j). Orbital interaction energy, also known as the Fock matrix element, is a reflection of orbital overlap between orbitals i and j. For the purposes of this work, the donor orbital, j, will always be the lone pair orbitals on O_γ, and the acceptor orbital, i, is the $\sigma^*(O_{3\beta}—P_\gamma)$ antibonding orbital in the methyl triphosphate model of ATP (Figure 10).

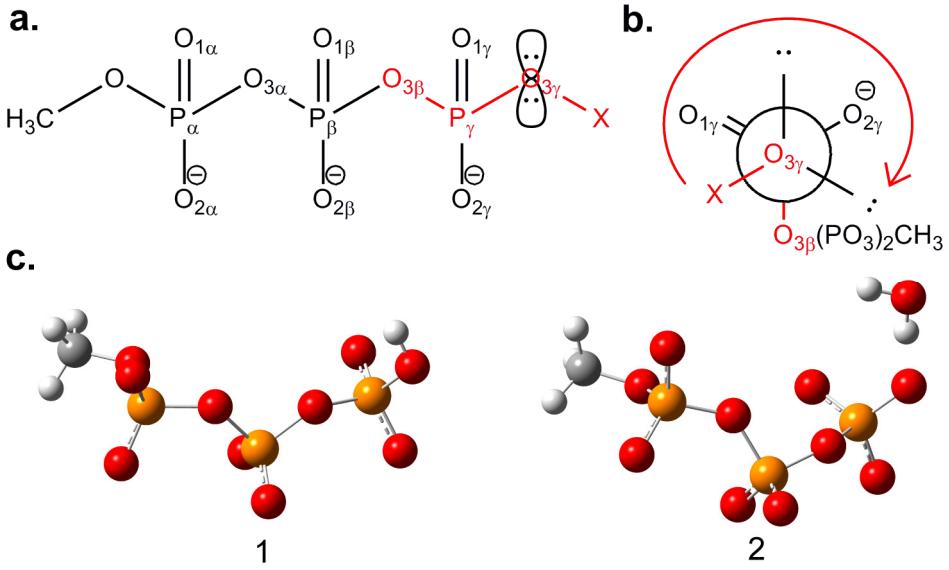


Figure 10. Orientation of lone pairs. a) Lewis structure of methyl phosphate triester. b) Corresponding Newman projection. The $O_{3\beta}—P_\gamma—O_{3\gamma}—X$ dihedral (highlighted in red) was rotated from -60° to 60° to change the orientation of lone pairs with respect to the $\sigma^*(O_{3\beta}—P_\gamma)$ anti-bonding orbital. c) X represents either a proton (structure 1) or a hydrogen bond donated by water (structure 2).

Methyl triphosphate has previously been employed as a simplified model for the study of hyperconjugative interactions in ATP.[15,109] Structures 1 and 2 were used to examine the effects of lone pair position on anomeric energy and $O_{3\beta}—P_\gamma$ bond length through rotation of the neighboring dihedral angles (Figure 10, shown in red). As each

dihedral angle ($\angle O_{3\beta}—P_{\gamma}—O_{3\gamma}—X$) was scanned, all other parameters of the structure were optimized, except in structure 2 where the $O_{3\gamma}•••OH_2$ hydrogen bond length was held constant at a typical 2.6 Å.[110] For structure 1, where one of the γ-oxygens is protonated, the $O_{3\beta}—P_{\gamma}—O_{3\gamma}—H$ dihedral angle was rotated from -60° to 60° in 10° increments. The $O_{3\beta}—P_{\gamma}—O_{3\gamma}•••OH_2$ dihedral angle in structure 2, was rotated from -60° to 60° in 20° increments. Changes in bond length, E(2) energy, interaction energy and orbital energy gap were calculated with reference to the values calculated for the staggered rotamers ($\angle O_{3\beta}—P_{\gamma}—O_{3\gamma}—X = 180°$). Throughout the text, lone pairs are labeled 1 or 2 to make referencing the lone pairs as they rotate around the $O_{3\beta}—P_{\gamma}—O_{3\gamma}—X$ dihedral angle more convenient. Only lone pairs that have $n(O_{\gamma}) \rightarrow \sigma^*(O_{3\beta}—P_{\gamma})$ E(2) energies that are greater than the NBO program default cutoff of 0.5 kcal/mol are discussed in this study.

The effects of different hydrogen bonding configurations were explored with structures 3 through 10 (Figure 11), which were built using GaussView5.0. Structures were then optimized while constraining the $O_{3\beta}—P_{\gamma}—O_{3\gamma}\cdots H(OH)$ and $O_{3\beta}—P_{\gamma}—O_{3\gamma}\cdots O(H_2)$ dihedral angles to fixed 180° or 90° values (both the oxygen and relevant hydrogen of water are fixed relative to $O_{3\beta}—P_{\gamma}$). The $O_{3\gamma}\cdots OH_2$ distance was again fixed at 2.6 Å and all other parameters were allowed to relax. All changes are measured in reference to structure 3, methyl triphosphate without explicit waters. The effects of hydrogen bond length were examined using structure 4. As the donor-acceptor distance was increased from 2.4 Å, other structural parameters were optimized, with the exception of the $O_{3\beta}—P_{\gamma}—O_{3\gamma}\cdots H(OH)$ and $O_{3\beta}—P_{\gamma}—O_{3\gamma}\cdots O(H_2)$ dihedral angles which were held at 180° in the *anti*-conformation.

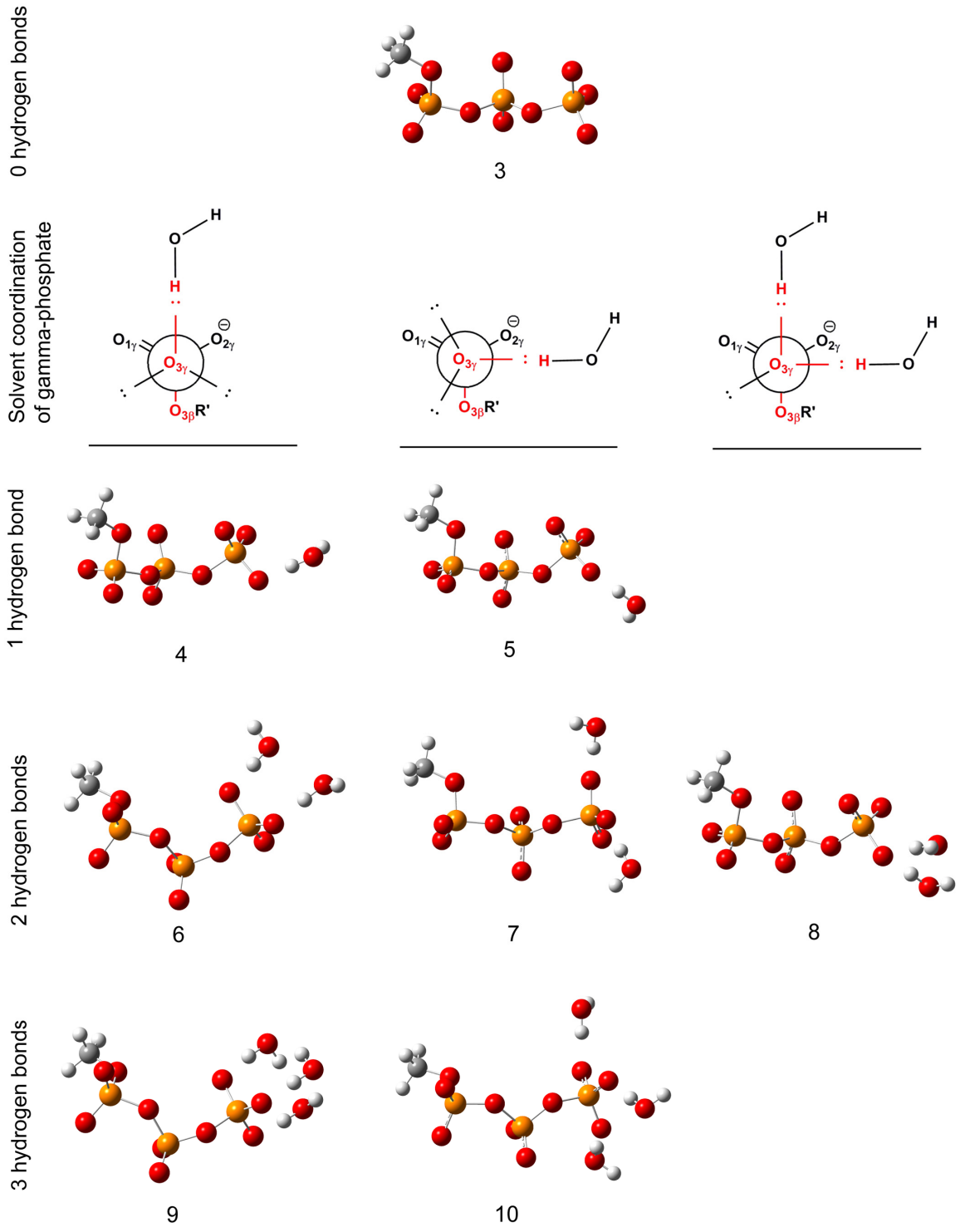


Figure 11. Eight structures with different hydrogen bonding configurations were used in a detailed examination of the effects of hydrogen bonding on phosphodiester bonds. Structure 3 has no hydrogen bonding interactions and serves as a reference. Waters are

placed at a $O_{3\beta}$—P_{γ}—O_{γ}•••OH_2 dihedral angle of 180° or 90° as depicted in red in the Newman projections above each column of structures. White = hydrogen, grey = carbon, red = oxygen, yellow = phosphorus.

RESULTS AND DISCUSSION

The effects of lone pair alignment on hyperconjugation and bond lengths in methyl triphosphate. – The anomeric effect is classically thought of as highly dependent on optimal orbital alignment of lone pair and antibonding orbitals.[16,111] Thus, the anomeric interaction, $n(O_{\gamma}) \rightarrow \sigma^*(O_{3\beta}—P_{\gamma})$ is expected to be diminished by rotation of $n(O_{\gamma})$ out of the anti-periplanar position. To characterize the influence of the anomeric effect on the $O_{3\beta}—P_{\gamma}$ bond in methyl triphosphate, orbital occupancies, E(2) energies and $O_{3\beta}—P_{\gamma}$ bond length were calculated as a function of torsional rotation of the $O_{3\beta}—P_{\gamma}—O_{3\gamma}—H$ dihedral angle in structure 1 (Figure 10). To understand the basis of variation in anomeric energy, E(2) energy was broken down into its components of interaction energy, $F_{i,j}$ and orbital energy gap, E_j-E_i (Equation 1), both as a function of the $O_{3\beta}—P_{\gamma}—O_{3\gamma}—H$ dihedral angle.

The anomeric energy is found to be highly correlated with the $O_{3\beta}—P_{\gamma}$ bond length ($R^2 = 0.92$, Figure 12). A parallel increase was observed in the *anti*-bonding orbital occupancy with a correlation between occupancy and bond length of $R^2 = 0.98$. If the changes in $O_{3\beta}—P_{\gamma}$ were due to inductive withdrawal into the $P_{\gamma}—O_{3\gamma}$ bond, decrease in the $O_{3\beta}—P_{\gamma}$ bonding orbital occupancy would be expected. However, the occupancy was independent of the $O_{3\beta}—P_{\gamma}—O_{3\gamma}—H$ rotation, suggesting that the observed change in bond length does not result from an inductive effect.

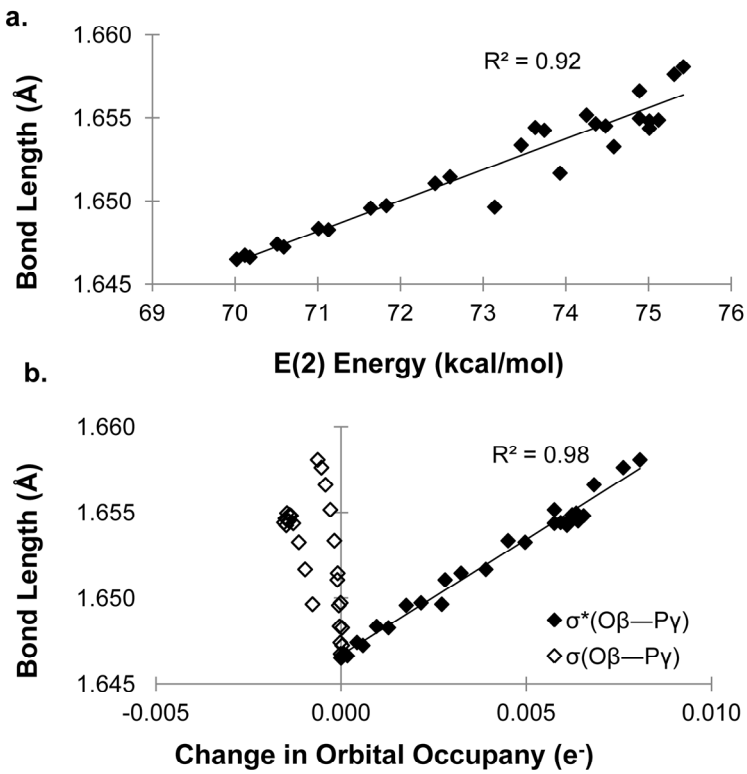


Figure 12. a) Calculated E(2) hyperconjugative energies and $O_{3\beta}$—P_γ bond lengths for rotamers of structure 1. Changes are measured with reference to ∠$O_{3\beta}$—P_γ—$O_{3\gamma}$—H = 180°. b) Correlation between $O_{3\beta}$—P_γ bond length and σ($O_{3\beta}$—P_γ) orbital occupancy.*

The anomeric effect is maximal when lone pair 2 is *anti*-periplanar to $O_{3\beta}$—P_γ and the proton is at a 90° dihedral angle relative to $O_{3\beta}$—P_γ (Figure 13). At this orientation of the hydroxyl group, lone pair 2 accounts for almost 100% of the total E(2) with σ*($O_{3\beta}$—P_γ) (Figure 13a, blue line). The, the overall change in E(2) energy (Figure 13a, black line) with angular rotation occurs because lone pair 2 is a much stronger electron density donor than both lone pair 1 and σ(O—H) combined. Lone pair 2 is readily able to donate electron density because of its smaller orbital energy gap with the σ*($O_{3\beta}$—P_γ) orbital (ca. 600 kcal.mol) as compared to the gap between σ*($O_{3\beta}$—P_γ) and lone pair 1 or between σ*($O_{3\beta}$—P_γ) and σ($O_{3\gamma}$—H) (both ca. 900 kcal/mol, Figure 13c). Furthermore,

the maximum orbital overlap between $\sigma^*(O_{3\beta}—P_\gamma)$ and lone pair 2 is more than twice the maximum orbital overlap between $\sigma^*(O_{3\beta}—P_\gamma)$ and lone pair 1 or between $\sigma^*(O_{3\beta}—P_\gamma)$ and $\sigma(O_{3\gamma}—H)$. Together the relatively lower orbital energy gap and higher interaction energy of lone pair 2 with $\sigma^*(O_{3\beta}—P_\gamma)$ dictate a 6 kcal/mol increase in E(2) energy at $\angle O_{3\beta}—P_\gamma—O_{3\gamma}—H = 180°$ over $\angle O_{3\beta}—P_\gamma—O_{3\gamma}—H = 90°$. This change in E(2) energy is mirrored by a 0.012 Å increase in $O_{3\beta}—P_\gamma$ bond length.

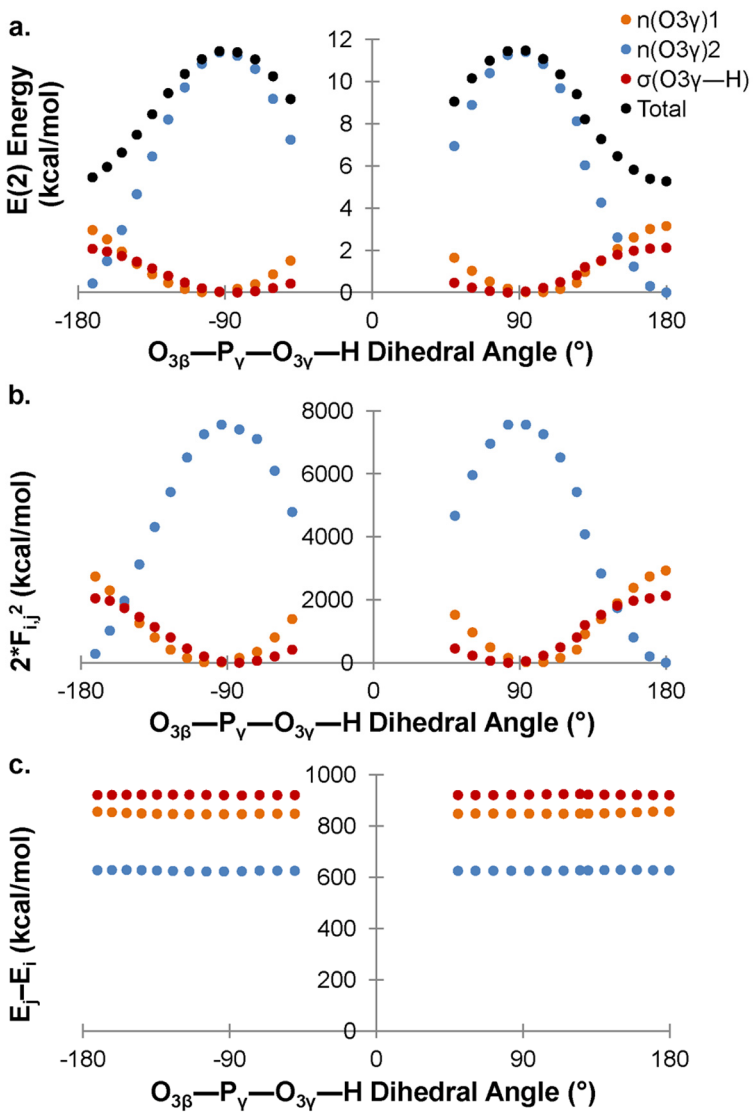


Figure 13. a) E(2) energy shows a strong angular dependence for rotation of the $O_{3\beta}$—P_γ—$O_{3\gamma}$—H dihedral angle. b) The orbital overlap, as mirrored in the numerator of equation 1, $2F_{ij}^2$ and c) the energy gap for the hyperconjugative interaction with $\sigma^*(O_{3\beta}$—$P_\gamma)$ and lone pairs n1($O_{3\gamma}$) (orange), n2($O_{3\gamma}$) (blue) or σ(O—H) (red).

Consistent with previous reports, our study demonstrates that hypercongugative n(O) → σ*(O—P) interactions are prominent in the lengthening / weakening of vicinal $O_{3\beta}$—P_γ bonds. There is indeed an anomeric component to the interaction, the $O_{3\beta}$—P_γ

bond length changing 0.01 Å, dependent on the orientation of the donating lone pair relative to the *anti*-bonding orbital. Unique to this study is that changes in hyperconjugation, while dependent on the orbital energy gaps, are most responsive to changes in interaction energy, as mirrored in the numerator of Equation 1, $2F_{ij}^2$.

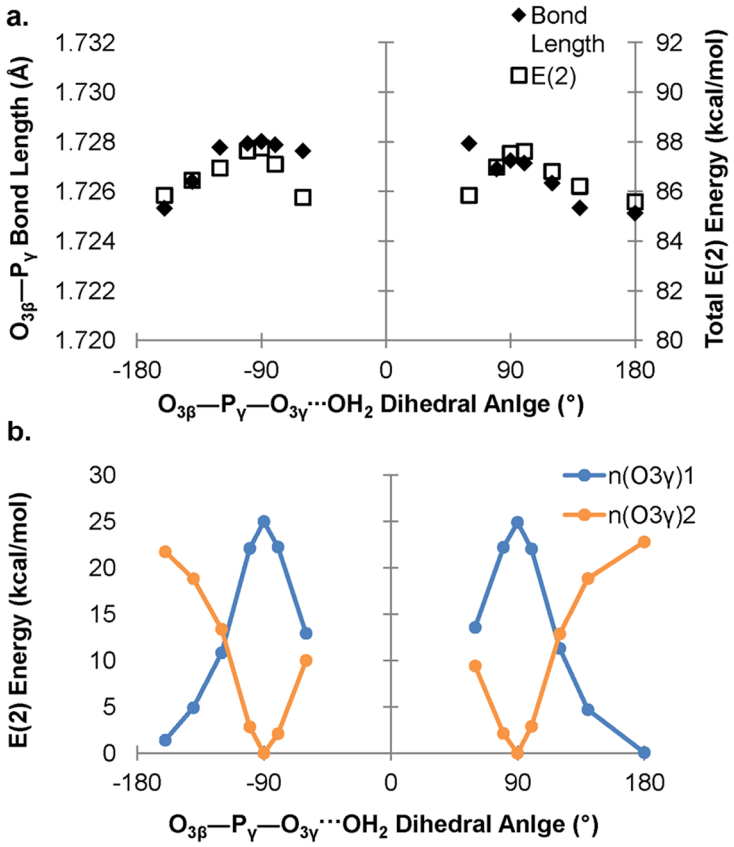


Figure 14. Orientation of a hydrogen bond from water relative the $O_{3\beta}$—P_γ bond has little impact on $O_{3\beta}$—P_γ bond length. a) Neither the $O_{3\beta}$—P_γ bond length or E(2) energy vary greatly as a function of $O_{3\beta}$—P_γ—$O_{3\gamma}$•••OH_2 dihedral angle. The maximal change in bond length is only 0.003 Å and in E(2) energy is only 2 kcal/mol. b) When one lone pair orbital rotates away from an optimal anti-periplanar orientation, the other improves. The E(2) energies are nearly symmetric, in spite of the fact that a water is interacting with lone pair 2 (orange), but not 1 (blue).

Hyperconjugative bond lengthening is independent of solvation geometry. – Hyperconjugative interactions have previously been calculated for dihydrogen- and metaphosphate interactions with waters, where there is strong donation of electron density from the dihydrogen-or metaphosphate oxygen lone pairs into the σ*(O—H) *anti*-bonding orbital of water.[112] Thus, it might be expected that water would withdraw electron density from the lone pairs in methyl triphosphate, thus weakening the $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$ interaction. Indeed addition of a hydrogen bond between water and $O_{3\gamma}$ of methyl triphosphate (structure 4 in Figure 11) is accompanied by a 0.017 Å decrease on $O_{3\beta}—P_\gamma$ bond length and a 4.8 kcal/mol reduction in E(2) energy. As a 0.01 Å change in bond length was induced in structure 1 through rotation of the OPOH dihedral angle, it was hypothesized that analogous rotation of the water would induce corresponding changes in bond length. To test this, the $O_{3\beta}—P_\gamma—O_{3\gamma}•••OH_2$ dihedral angle in structure 2 (Figure 10) was rotated in increments of 20° from -60° to +60° (see Methods).

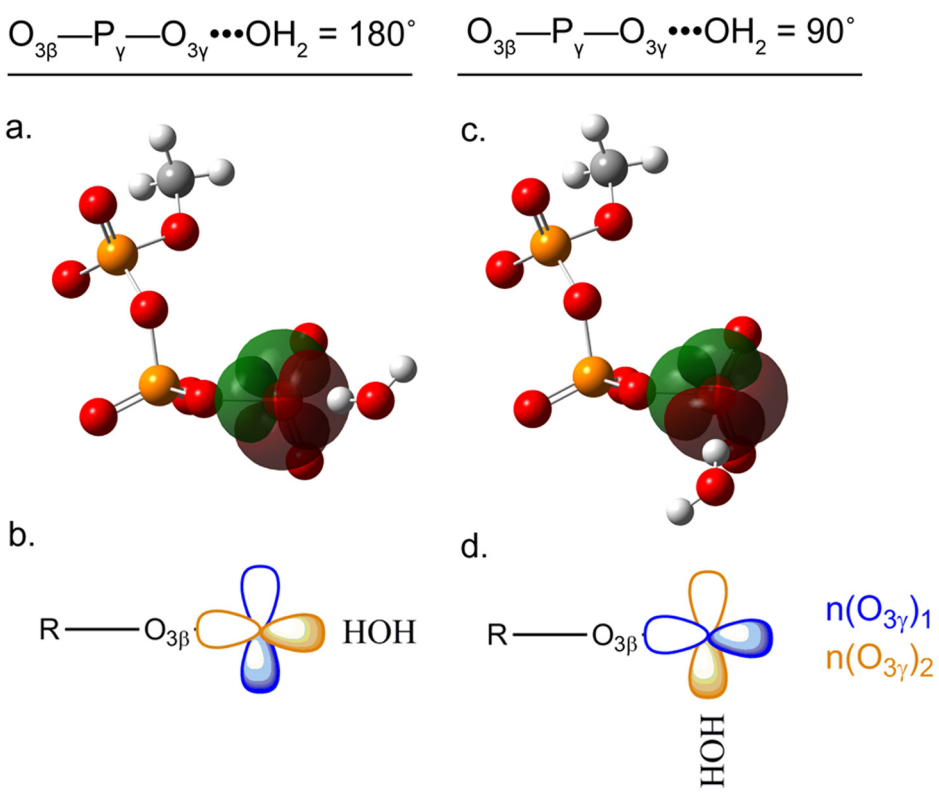


Figure 15. Images of the $O_{3\gamma}$ lone pairs 1 and 2 from Figure 14b at a) $O_{3\beta}$—P_γ—$O_{3\gamma}$•••OH_2 = 180° and b) $O_{3\beta}$—P_γ—$O_{3\gamma}$•••OH_2 = 90°. Water is always coordinated to lone pair 2 (colored orange in parts b and d). In a) water interacts with lone pair 2, which in this configuration, is anti-periplanar to $O_{3\beta}$—P_γ whereas in b) water interacts with lone pair 2 which is at a 90° angle to $O_{3\beta}$—P_γ. When lone pair 2 is at 90°, it contributes nothing to the overall E(2) energy of $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$. However, the E(2) energy of $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$ and the $O_{3\beta}$—P_γ bond length are similar in both orientations. Atoms are colored; white (hydrogen), grey (carbon), yellow (phosphorus) and red (oxygen). The orbital lobes of the lone pairs on $O_{3\gamma}$ are colored red and green, depicting opposite polarities.

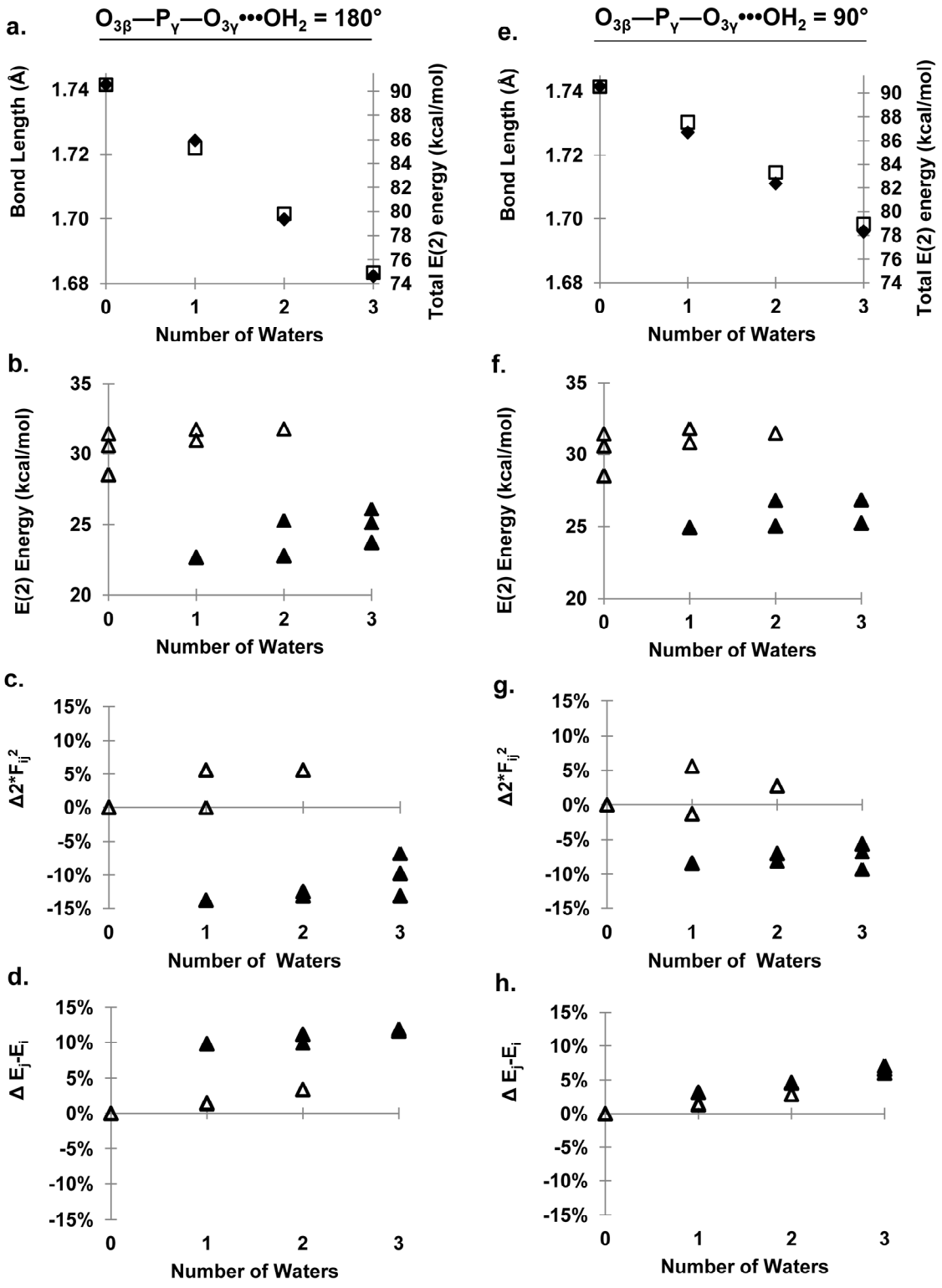


Figure 16. Waters coordinated at 180° (a - d) or 90° (e - f) from the $O_{3\beta}$—P_γ bond both decrease (a, e) $O_{3\beta}$—P_γ bond length and (b, f) hyperconjugation between γ-oxygen lone

pairs and the σ($O_{3\beta}$—P_γ) anti-bonding orbital. Open triangles represent values of hyperconjugative interactions where there is no hydrogen bond. Closed triangles represent values where there is a hydrogen bond. (c, g) In both orientations, water reduces $2F_{ij}^2$ relative to structure 3, water-free methyl triphosphate. (d, h) However only structures with waters at 180° show a pronounced change in orbital energy gap. This discrepancy accounts for the slightly larger depression in $O_{3\beta}$—P_γ bond length for structures with waters at 180°. Only orbitals in the anti-periplanar orientation contribute to E(2) energy and thus values for the other two orbitals on each oxygen were not plotted.*

Solvent hydrogen bonds oriented at 90° or 180° with respect to $O_{3\beta}$—P_γ induce similar perturbations in $O_{3\beta}$—P_γ bond lengths, differing by only 0.003 Å (Figure 14). Likewise total E(2) interaction energies are similar (E(2) for $n(O_{3\gamma}) \rightarrow \sigma^*(O_{3\beta}$—$P_\gamma)$ differing by 2.2 kcal/mol). Thus, the orientational dependence of the hydrogen-bonded system is about one third that of the protonated system. In the solvated system, each of the two higher energy $O_{3\gamma}$ lone pair orbitals can have a similar E(2) interaction if *anti*-periplanar to $\sigma^*(O_{3\beta}$—$P_\gamma)$, as shown in Figure 14b. With dihedral rotation of the lone pairs about $O_{3\beta}$—P_γ (as they follow the scanned direction of the water hydrogen bond), the change in one E(2), as the orbital rotates away from an optimal *anti*-periplanar orientation, is compensated approximately by the change in the other as it becomes more *anti*-periplanar. Which of these lone pairs is participating in a hydrogen bond has relatively little impact, because, as discussed below, both lone pair orbital interactions are affected by the presence of the water simultaneously, without much directional specificity, and to a degree that had not been anticipated.

Water interactions can decrease orbital overlap of additional lone pairs with σ(O3β—Pγ).* – The difference in orbital interactions are shown in Figure 15. When $\angle O_{3\beta}—P_{\gamma}—O_{3\gamma}\bullet\bullet\bullet OH_2 = 90°$, it is the non-hydrogen bonded $n(O_{3\gamma})_1$ that contributes 100% of the $O_{3\gamma}$ E(2) energy (Figure 14). As there is not a direct interaction between $n(O_{3\gamma})_1$ and the hydrogen bond from water (Figure 15c), it was surprising that the E(2) energy was affected by the presence of the water. This led to a broader analysis of the electronic differences between conformers with $\angle O_{3\beta}—P_{\gamma}—O_{\gamma}\bullet\bullet\bullet OH_2 = 180°$ and 90°. Analysis started with a study of E(2) energy as a function of the number of O_{γ} hydrogen bonds (Figure 16) for structures 3 through 7, 9 and 10 (Figure 11).

In Figure 16, $O_{3\beta}—P_{\gamma}$ bond length, E(2) energy, and its components $2F_{ij}^2$ and E_j-E_i of the of $n(O_{\gamma}) \rightarrow \sigma^*(O_{3\beta}—P_{\gamma})$ hyperconjugative interaction are plotted against number of waters coordinated to γ-oxygen lone pairs. Figure 16a-d represents structures 4, 6 and 9 (Figure 11) where the water is hydrogen bonded to the lone pair that is *anti*-periplanar to $O_{3\beta}—P_{\gamma}$ (Figure 15a).

Figure 16e-f represents structures 5, 7 and 10 where the water is bound to the lone pair that is offset by 90° and is thus not participating in the $n(O_{\gamma}) \rightarrow (O_{3\beta}—P_{\gamma})$ hyperconjugative interaction (Figure 15c). There are only small differences in $O_{3\beta}—P_{\gamma}$ bond length and E(2) values for structures with hydrogen bond(s) oriented at $\angle O_{3\beta}—P_{\gamma}—O_{\gamma}\bullet\bullet\bullet OH_2 = 180°$ (Figure 16a) versus $\angle O_{3\beta}—P_{\gamma}—O_{\gamma}\bullet\bullet\bullet OH_2 = 90°$ (Figure 16e). For example, a hydrogen bond with the *anti*-periplanar lone pair relative to the $O_{3\beta}—P_{\gamma}$ bond induces a decrease in the $O_{3\beta}—P_{\gamma}$ bond length by about 0.02 Å, and with three waters, $O_{3\beta}$-P_{γ} is shortened by about 0.06 Å. When the hydrogen bond is offset 90°, these values are only slightly less (0.015 Å and 0.04 Å respectively). Similarly the

addition of a hydrogen bond at $\angle O_{3\beta}—P_\gamma—O_\gamma•••OH_2 = 180°$ induces a decrease in E(2) energy of about 6 kcal/mol (Figure 16b). If the hydrogen bond is at $\angle O_{3\beta}—P_\gamma—O_\gamma•••OH_2 = 90°$ the reduction in E(2) is closer to 4 kcal/mol (Figure 16f). Particularly striking is the relative similarity of plots c and g in Figure 16. It had been expected that a water would only affect the orbital with which it is hydrogen-bonded. However we find that the $2F_{ij}^2$ terms, are decreased for both orbitals 1 and 2 regardless of hydrogen bond orientation by roughly 10%.

In a first (unsuccessful) attempt to rationalize direction-independent changes to F_{ij}^2, we examined how hydrogen-bonding might affect hybridization of orbitals not directly participating in the hydrogen bond. A lone pair starting with 99.7% p-character gains 8.9% s-character on hydrogen bonding, but other lone pairs are affected negligibly (< 0.02%). Thus, direction-independent effects do not appear to result from widespread changes in hybridization.

Hydrogen bonding lengthens the non-bridging $P_\gamma—O_\gamma$ bond of the coordinated oxygen (Figure 17). An increase in the $P_\gamma—O_\gamma$ bond length physically increases the distance between the lone pair orbitals on O_γ and the $\sigma^*(O_{3\beta}—P_\gamma)$ *anti*-bonding orbital (Figure 17, inset) thus decreasing their overlap. This explains computed changes to F_{ij}^2 that are independent of the direction of hydrogen bonding, because all hydrogen bonds to the O_γ increase the $P_\gamma—O_\gamma$ bond length. Consistent with this rationalization, there is a strong correlation between $\Delta 2F_{ij}^2$ and $P_\gamma—O_\gamma$ bond lengths for structures 4 through 10 ($R^2 = 0.90$).

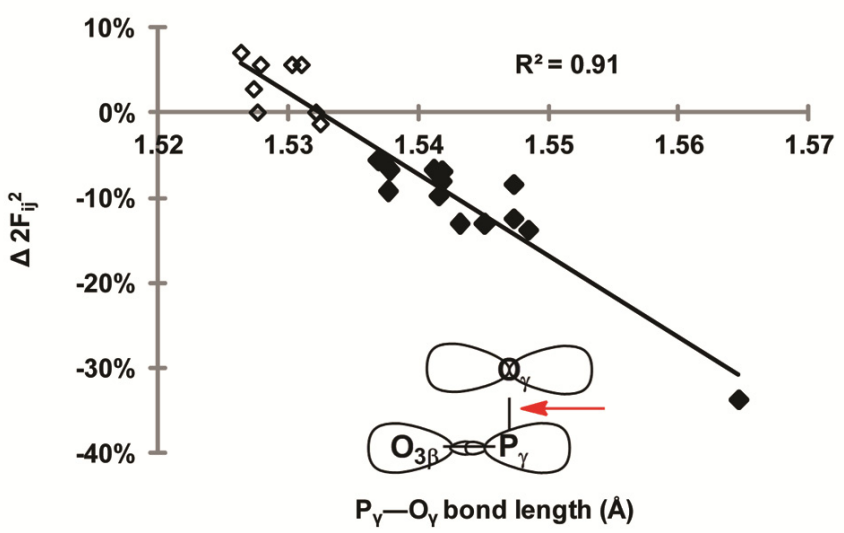


Figure 17. For structures 4 through 10, there is a strong correlation between $\Delta 2F_{ij}^2$ of the interaction $n(O_\gamma) \rightarrow (O_{3\beta}—P_\gamma)$ and non-bridging $P_\gamma—O_\gamma$ bond length. Changes are measured relative to structure 3. Open diamonds represent $P_\gamma—O_\gamma$ bond lengths where there is no hydrogen bond with O_γ. Closed diamonds represent $P_\gamma—O_\gamma$ bond lengths where there is a hydrogen bond with O_γ. Inset: the schematic highlights (red arrow) the bond that is lengthened indirectly by any configuration of hydrogen bond to O_γ, thereby affecting the proximity and overlap of the O_γ lone pair orbitals and the $\sigma^*(O_{3\beta}—P_\gamma)$ anti-bonding orbital.

Two hydrogen bonds with a single oxygen has the same effect as two hydrogen bonds with two different oxygens. – The change in $O_{3\gamma}—P_\gamma$ bond length is linearly proportional to the number of hyperconjugation-modulating solvent hydrogen bonds on different oxygens (Figure 16). It has also been shown that changes in the $n(O_\gamma) \rightarrow (O_{3\beta}—P_\gamma)$ hyperconjugative interaction are not greatly dependent on the direction of hydrogen bonding. Thus, the effects on E(2) energy and bond length of two hydrogen bonds on the same oxygen, one at 180° and one at 90°, might be expected to be additive.

This postulate was examined by comparing structures 6, 7 & 8 with structure 3. Unlike the desolvated structure 3, structures 6, 7 & 8 are coordinated by two waters. In structures 6 & 7, the waters are coordinated to different oxygens, but in structure 8, both waters interact with the same γ-oxygen. The $O_{3\beta}$—P_γ bond lengths (and relevant E(2) energies) of structure 6, 7 and 8 are very similar to each other but different from structure 3 (Table 1). Thus, whether interacting with the same or different γ-phosphate oxygens, the effects of multiple hydrogen bonds are, to a first approximation, additive and independent of orientation relative to the σ* orbital. This is of potential enzymological significance, because an oxygen of an ATP bound in an active site can have multiple hydrogen-bonding interactions.[37,43,113,114] The effects of all upon catalysis will need to be integrated, not just interactions with lone pairs *anti*-periplanar configuration to σ*.

Table 1. $O_{3\beta}$—P_γ Bond lengths and $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$ E(2) energies of structures 3, 6, 7 and 8.

Structure	$O_{3\beta}$—P_γ Bond length	Total E(2) of $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$
3	1.742	90.6
6	1.700	79.8
7	1.711	83.3
8	1.711	81.9

Dependence on the Strength of the Hydrogen Bond. – The strength of the hydrogen bond interaction was modulated by changing the constrained hydrogen bond

length in structure 4. With a hydrogen bond length of 2.4 Å, the $O_{3\beta}$—P_{γ} bond is 1.718 Å, 0.024 Å shorter than without a water molecule (structure 3). The calculated change in phosphodiester bond energy due to loss of a 2.4 Å-hydrogen bond is approximately 21 kcal/mol. (Thus the loss of two hydrogen bonds would be commensurate with the ~40 kcal/mol empirically measured activation barriers of representative phosphoryl transfer enzyme reactions)[115-119] The $O_{3\beta}$—P_{γ} bond length increases by ~ 0.02 Å per Å increase in the hydrogen bond length (Figure 18). At a hydrogen-acceptor distance of 3.4 Å, the hydrogen bond has negligible impact upon the $O_{3\beta}$—P_{γ} bond length and E(2) energy of $n(O_{\gamma}) \rightarrow \sigma^*(O_{3\beta}—P_{\gamma})$ interaction. Hydrogen bonds are usually considered to have both electronic and classical electrostatic components.[120,121] The 3.4 Å limit on modulation of the ATP stereoelectronics is consistent with the understanding that the covalent component of a hydrogen bond decreases with distance.

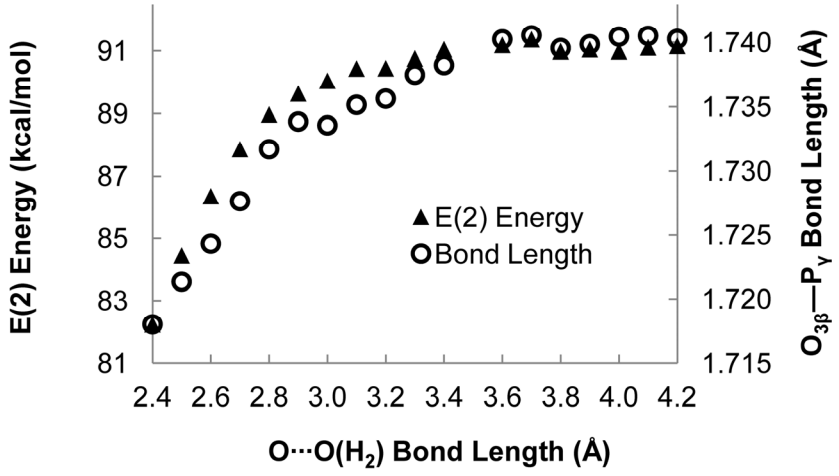


Figure 18. Effect of hydrogen bond length, $O_{\gamma}\cdots H(OH)$ on $O_{3\beta}$—P_{γ} bond length for structure 8.

CONCLUSION

It is increasingly clear that interactions between lone pair orbitals and *anti*-bonding orbitals are central to the properties of several high energy molecules of biochemical importance. Earlier characterizations focused on anomeric effects where the lone pair orbitals were optimally aligned for interaction with the σ* acceptor orbital.

Here, it is shown that considerable conformational variation can be accommodated: the effects of coordinating to the *anti*-periplanar and *gauche* lone pairs are similar in their reduction of the E(2) anomeric interaction and in decreasing the length of the vicinal phosphoanhydride bond. Interactions with the lone pairs modulate the anomeric effect largely through lengthening of non-bridging P—O bonds which changes the overlap between all lone pair orbitals on the coordinated oxygen and the σ* *anti*-bonding orbital, and this, in turn, affects the Fock interaction energy.

The significance of this finding is that a larger set of enzyme-substrate and solvent-substrate interactions needs to be considered in assessing environmental modulation of phosphoanhydride bond lability. Coordination of lone pairs, not only in the *anti*-periplanar direction, but also in *gauche* directions, has nearly the same potential to lessen an anomeric effect and conjointly stabilize the vicinal phosphoanhydride. Loss of such interactions by desolvation, as a substrate binds in an active site, could be one of the means with which an enzyme could destabilize the scissile bond in phosphoryl transfer reactions. It is also one of the ways that an enzyme could modulate the relative stability of the different phosphoanhydride bonds in a nucleotide, potentially impacting product specificity. Thus the investigation here lays a foundation for improving our

mechanistic understanding of the kinases and phosphatases that play critical roles in energy metabolism, cellular motility and signaling.

ACKNOWLEDGMENT

This work was supported in part by NIH R01GM77643 (MSC) and by pre-doctoral fellowships from the American Heart Association (Pacific Mountain Affiliate, predoctoral fellowship 09PRE2020112) and Vertex Pharmaceuticals Inc. (JCS). The authors thank Dr. Omar Davulcu for critical reading of the manuscript.

2. ENZYMES PROMOTE NUCLEOTIDE PHOSPHORYL TRANSFER THROUGH INTERACTIONS WITH THE BRIDGING OXYGEN THAT ENHANCE HYPERCONJUGATION

ABSTRACT

Phosphoryl transfer reactions are ubiquitous in nature, yet the chemical mechanisms remain in debate. In this work, analysis of Protein Data Bank crystal structures reveals motifs of substrate interactions that are widespread throughout many classes of enzyme. Specifically, it is found that hydrogen bonds with the bridging oxygen of the scissile phosphoanhydride bond in nucleotide substrates are prevalent. The importance of such interactions is assessed experimentally in the arginine kinase model system. Mutation of Arg_{280}, which forms a hydrogen bond with the $O_{3\beta}$ of ATP, is found to decrease k_{cat} 100-fold, with little change to K_M. Ground state quantum mechanical (QM) modeling shows that hydrogen bonding with the bridging oxygen induces scissile bond elongation by enhancing hyperconjugation in the ligand. Other hydrogen bonds that calculation predicts would decrease hyperconjugation are correspondingly rare in the structural database. Hyperconjugation provides a unifying rationalization not only for interactions identified from the structural database, but it also explains the changed bond order previously inferred from kinetic isotope and infrared experiments. Both the experimental kinetics and QM modeling lend validation to our top-down statistical analysis and its ability to uncover unifying mechanistic principles from large samples of structures at a level of detail beyond the precision of an individual structure. The current application of this approach shows that modulation of stereoelectronics is a component

contributing to the catalysis of a broad spectrum of phosphoryl transfer enzymes with wide diversity in proposed mechanisms.

AUTHOR SUMMARY

Catalysis of phosphoryl transfer is critical for biology, but few unifying principles have emerged in enzyme mechanisms. We start by analyzing enzyme-substrate interactions in a database of protein crystal structures. We find that hydrogen bonds with the bridging oxygen of the cleavable bond in the substrate are prevalent in phosphoryl transfer enzymes. Quantum mechanical analysis shows that these hydrogen bonds weaken the bond to be cleaved by increasing substrate hyperconjugation. Experimental enzyme kinetics with Arginine kinase active site mutants confirms that these interactions have a significant impact upon catalytic rate. The ubiquity of this type of interaction indicates an element of mechanism that is common to a diverse array of kinases and phosphatases.

INTRODUCTION

Enzymes that catalyze the transfer of a phosphate from ATP are widespread in biology. Free energy liberated in hydrolysis of phosphoandyride bonds, such as those in ATP, is essential for cellular energy metabolism, motility and the generation of transmembrane potentials. Much has been learned about enzyme-catalyzed phosphoryl transfer through the detailed study of individual enzymes. The most frequently cited mechanisms include; a) positioning of substrates for phosphate transfer; b) base-assisted activation of the nucleophile; c) activation of the electrophile; and finally d) electrostatic stabilization of the transition state.[27] Additional mechanisms have been proposed for specific enzymes, including: proton relay systems,[43] substrate-assisted catalysis [42] and

strain of the β and γ phosphate groups.[27] Phosphoryl transfer enzymes are a broad class within which several mechanisms may have evolved. However, one is struck by the diversity in mechanistic proposals and the lack of consensus on elements of mechanism that might be shared between enzyme families.

The Protein Data Bank (PDB) provides an opportunity for a top-down analysis of common active site configurations. It was hoped that this could complement traditional reductionist investigations of individual systems. Three classes of structures are compared here: those that cleave an $O_{3\beta}$—P_{γ} bond (Figure 19a); those that cleave a P_{α}—$O_{3\alpha}$ bond (Figure 19b), and a "control" group of those that bind nucleotide without catalyzing phosphoryl transfer (Figure 19c). We then analyze systematic differences in the predominant interactions, gaining insights from the population trends that might be obscured by experimental error or other variation in individual structures. Limitations of the approach were clear from the start. Ground state structures predominate with limited information about transition states. Experimental errors in individual structures can be appreciable. Notwithstanding such limitations, with increasing database size, it was an opportune time to test our primary goal of whether additional functional insights could be revealed by a top-down analysis in which the errors of individual structures would be averaged out. We then set out to validate the primary inference from our database study through experimental kinetic analysis of site-directed mutants of the arginine kinase model system. Finally using quantum mechanical modeling we provide a rationale for how the highlighted interactions could have a mechanistic role in catalysis.

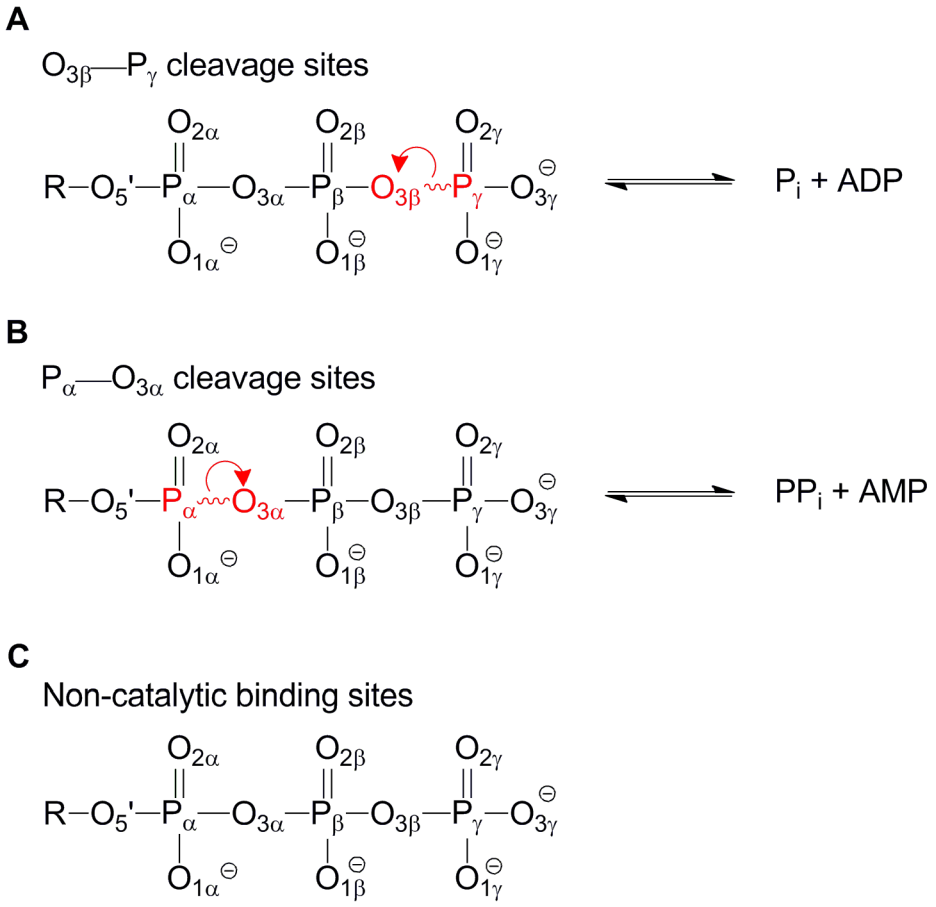


Figure 19. Enzyme structures can be categorized according to the fate of the bound nucleoside triphosphate (NTP). a) Phosphoryl transfer in which the $O_{3\beta}$—P_γ bond is cleaved. b) Reactions in which the P_α—$O_{3\alpha}$ bond is cleaved and c) Structures where the bound NTP does not undergo a chemical reaction. Red lettering indicates the atoms in the scissile bond and red arrows depict the transfer of electrons in going from reactants to products.

MATERIALS AND METHODS

Survey of Protein Data Bank crystal structures – Structures of protein-nucleotide complexes (Table S1) were downloaded from the PDB, including those containing ATP, UTP, TTP, GTP, CTP or their analogs. Two sets of data were compiled, one containing structures up to 2.0 Å resolution, where solvent molecules were defined with confidence,

and a larger group (inclusive of the first) up to 2.7 Å resolution from which solvent water was excluded, due to insufficient confidence in their positions at this resolution. Of the 1866 protein-NTP complexes to 2.7 Å, as of October 3rd, 2012, 305 were non-redundant and satisfied the inclusion criteria as detailed in the supplemental material (Table S2 and Table S3). Grounds for exclusion included adoption of an open-state conformation known to be non-catalytic. Of the 305 vetted structures, 134 were at resolutions higher than 2.0 Å. Molprobity's Reduce [122] was used to flip asparagine, glutamine and histidine side chains for optimal hydrogen bonding.

Structures were grouped by the site of bond cleavage in the ligand: 1) at the $O_{3\beta}$—P_{γ} bond (Figure 19a); 2) at the P_{α}—$O_{3\alpha}$ (Figure 19b); and 3) those in which the ligand binding site is non-catalytic (Figure 19c). The groups had 155, 100, and 50 PDB structures in the 2.7 Å set, respectively and 72, 45 and 17 structures in the 2.0 Å set. An underlying premise was that enzymes might share catalytic features regardless of base-type (GTP, ATP, *etc*.). The set of proteins with non-catalytic binding sites serves as a control to distinguish potentially catalytic interactions from those primarily involved in ligand-binding. Comparisons between the sets of enzymes with P_{α}—$O_{3\alpha}$ and $O_{3\beta}$—P_{γ} cleavage sites also help to distinguish binding and catalytic interactions, due to spatial separation of binding and catalytic sub-sites.

Interactions were identified after expanding the crystallographic symmetry and adding riding hydrogen atoms. These riding hydrogens were not seen directly by x-ray crystallography, but their positions were inferred from the heavy atom coordinates. Assessed interactions included hydrogen bonds between enzyme and ligand or between ligands and coordinated metal ions, all analyzed using an in-house Python program.

Heavy atom distance and angle criteria were based on previous work [123]: O•••D distance ≤ 3.5 Å, P—O•••D angle ≥ 90°, where D is the hydrogen bond donor, P is the ligand phosphorus atom, and O is the oxygen atom of the ligand. An additional criterion was added: O•••D—R angle ≥ 90° where R is the antecedent atom of a hydrogen bond donor. These criteria depend only on the heavy atoms directly observable and can therefore be applied directly to medium resolution crystal structures. Assessments were also made using additional criteria possible once riding hydrogens were added: O•••H ≤ 2.7 Å, P—O•••H ≥ 80°and O•••H—D ≥ 80°. Inspection revealed atom pairs satisfying the above criteria that were actually atoms coordinating a common metal ion. These pairs were excluded using criteria given in Figure S1. Metal interactions were subject only to a distance cut-off of 2.8 Å derived from analysis of metal coordination in the Cambridge Structural Database.

Ligand oxygens were organized into six groups: Non-bridging γ-, β- and α-oxygens, and bridging β-, α- and O5'-oxygens. Distributions of enzyme-ligand interactions for each of the groups were compared using pair-wise two-tailed t-tests in RStudio with pooled standard deviations and the Benjamini & Hochberg correction for multiple comparisons. For each comparison, the null hypothesis was tested that the two sample means could be drawn from the same population, at a significance level of $\alpha = 0.05$. Results are fully tabulated in the supplemental material (Table S4 and Table S5).

Enzyme mutagenesis, purification and kinetic analysis – Mutants of *Limulus polyphemus* arginine kinase (AK) were made using the QuikChange mutagenesis kit (Stratagene). Previously published methods were used to purify AK and measure kinetic parameters.[115] Initial velocities were measured using a 6 x 6 grid of substrate

concentrations: ATP at 0.2, 0.4, 0.8, 1.2, 1.6, and 2.0 mM, and arginine at 0.2, 0.4, 0.8, 1.6, 3.2, and 7.0 mM. Duplicate measurements were made with separate protein preparations. Steady state parameters K_{ia}, K_m and V_{max} for a random order sequential bimolecular-bimolecular (bi-bi) reaction mechanism were fitted to the data by nonlinear least squares using SigmaPlot.

Quantum mechanical modeling of active site interactions – The Gaussian 09 program [95] was used for all electronic structures calculated with the B3LYP/6-311++G(d,p) level of theory [98] and an implicit bulk solvent modeled with the polarizable continuum model (PCM).[100] This level of theory had previously been shown to be appropriate for capturing perturbation effects.[15] Natural bond orbital (NBO) analysis was performed with the GenNBO5.0 program [104] using the B3LYP/6-311++G(d,p)-optimized structures and the HF/cc-PVTZ model chemistry.[105] The NBO program facilitates a Lewis-like description of electron density by transforming nonorthogonal atomic orbitals from a HF wave function into natural bond orbitals. Second order perturbation E(2) energy, given by Equation 1, estimates the magnitude of electron delocalization and hyperconjugation.

$$E(2) = q_i \frac{F_{i,j}^{\ 2}}{E_j - E_i} \quad \text{EQ. 1}$$

q_i is the donor orbital occupancy, and $E_j - E_i$ is the energy gap between donor orbital i and acceptor orbital, j. $F_{i,j}$ is the Fock matrix element which describes donor and acceptor orbital overlap. In this work, the donor will be a lone pair orbital of the γ-phophate oxygens and the acceptor orbital is $\sigma^*(O_{3\beta}—P_\gamma)$, the antibonding orbital of the $O_{3\beta}—P_\gamma$ bond in methyl triphosphate. Methyl triphosphate has been used previously, to model interactions in ATP (Figure 20).[124]

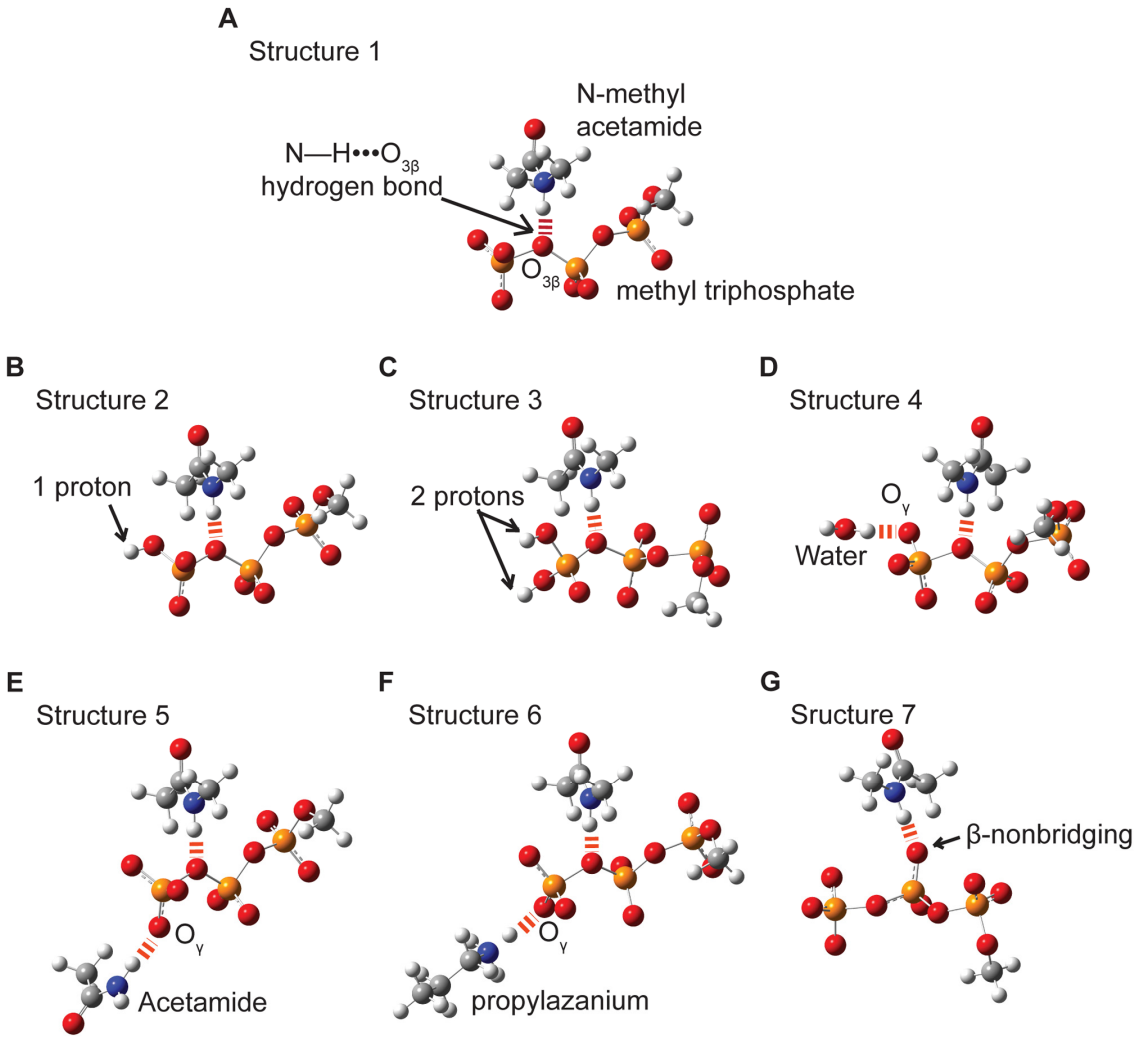


Figure 20. Models used to test the dependence of hyperconjugation and $O_{3\beta}$—P_γ bond length on enzyme-ligand interactions. N-methylacetamide (a) was used to model a (neutral) protein backbone amide hydrogen bond to $O_{3\beta}$, using methyl triphosphate to model an NTP nucleotide. Structures 2 (b) and 3 (c) were used to investigate the effects of protonation at the γ-oxygens. Additional active site hydrogen bonds, represented in Structures 4 through 6 (d-f), were used to assess the secondary effects of different types of O_γ hydrogen bonds. Acetamide (e) propylazanium (f) were used to model aspartate and lysine side chains respectively. Structure 7 (g) was used to investigate the impact of hydrogen bonding at a nonbridging β-oxygen on hyperconjugation and $O_{3\beta}$—P_γ bond

length. Hydrogen bonds are shown with dashed red lines. White = hydrogen, gray = carbon, red = oxygen, blue = nitrogen, orange = phosphorus.

The influence of a N—H•••$O_{3\beta}$ hydrogen bond on the $O_{3\beta}$—P_γ bond length and hyperconjugation was evaluated with Structures 1 through 6 (Figure 20a-f). Structure 7 was used to evaluate the effects of a hydrogen bond to a nonbridging β-oxygen (Figure 20g). N-methylacetamide was used to model a peptide backbone (neutral) hydrogen bond donor. The N—H•••O_β hydrogen bond was weakened by increasing its length from 2.5 Å to 3.5 Å in 0.1 Å increments for Structures 1 through 3 and in 0.2 Å increments for Structures 4 through 7. These changes were made by translating the N-methylacetamide, fixing the $O_{3\beta}$ and N atoms at a specific hydrogen bond distance, then optimizing other aspects of the structure with the N—H•••O_β angle fixed at 170°. Structures 2 through 6 were used to probe the electronic effects of either protontion or hydrogen bonding at the γ-oxygens of methyl triphosphate (Figure 20b-f). For protonated structures the $O_{3\beta}$—P_γ—O_γ—H dihedral angle was fixed at 180°. Acetamide was used as a model for an aspartate side chain in Structure 5 and propylazanium was used as a model for a lysine side chain in Structure 6. All hydrogen bonds to γ-oxygens, were fixed at a D•••O_γ distance of 2.8 Å, where D is the hydrogen bond donor heavy atom. The D•••O_γ—P_γ—$O_{3\beta}$ and the (D)H•••O_γ—P_γ—$O_{3\beta}$ dihedral angles were each fixed at 180°. Changes in $O_{3\beta}$—P_γ bond lengths and in E(2) energies were calculated relative to each molecule without a hydrogen bond between N-methylacetamide and O_β.

RESULTS

Database analysis of enzyme-nucleotide interactions – For the non-redundant protein-nucleotide complexes in the Protein Data Bank, interactions with nucleotide

oxygens were tabulated for each crystal structure. Assessed interactions include hydrogen bonds from active site amino acids, water, coordinated metal ions or other substrates. The numbers of interactions were compared between phosphoryl transfer enzymes (Figure 19a) and two control groups (Figure 19b, c) to distinguish potentially catalytic from non-catalytic associations. T-tests reveal two overall trends (Table S4 and Table S5), that interactions with both non-bridging β-oxygens and with the bridging oxygen of a scissile phosphoanhydride bond are more prevalent in phosphoanhydride bond cleaving sites than in non-catalytic nucleotide binding sites.

For non-bridging β-oxygens, the mean number of interactions in both $O_{3\beta}$—P_{γ} – cleaving sites (μ= 4.2, μ_{σ} = 0.2) and P_{α}—$O_{3\alpha}$–cleaving sites (μ = 3.3, μ_{σ} = 0.1) is significantly greater than in non-catalytic sites (μ= 2.5, μ_{σ} = 0.2; $p < 0.05$; 2.7 Å; Figure 21a, Table S4a).

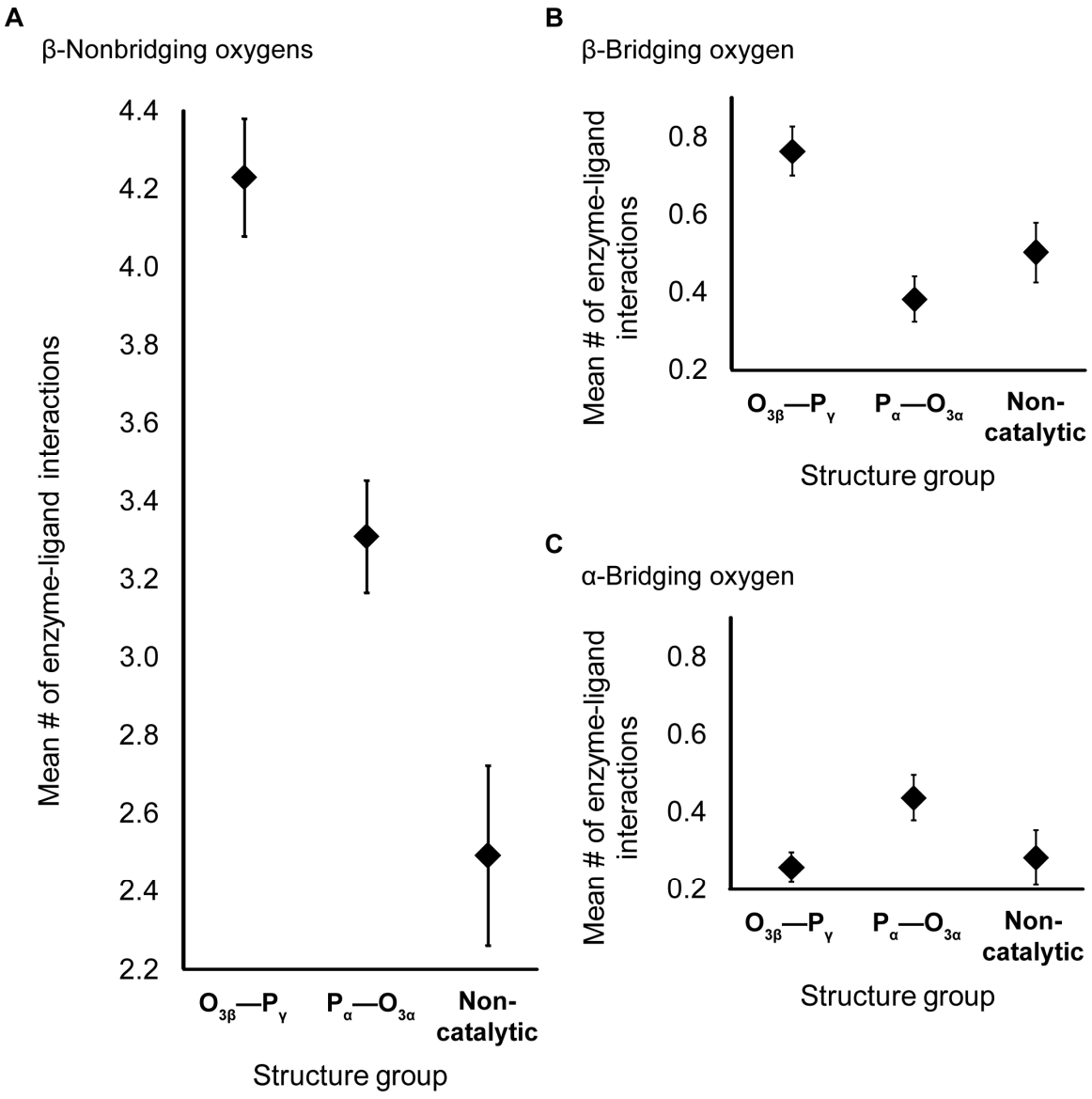


Figure 21. Mean number of enzyme-ligand interactions. Interactions are shown with nonbridging β- (a), bridging β- (b) and bridging α-oxygens (c) in the $O_{3\beta}$—P_γ cleaving-, P_α—$O_{3\alpha}$ cleaving- and non-catalytic-NTP-binding sites. The resolution cutoff of structures used is 2.7 Å, and water is not included. Bars show standard errors.

Although not every protein interacts with the bridging $O_{3\beta}$, interactions are more common in $O_{3\beta}$—P_γ cleaving sites (mean number of 0.8, $\mu_\sigma = 0.1$) than in either P_α—$O_{3\alpha}$ cleaving sites ($\mu = 0.4$, $\mu_\sigma = 0.1$) or non-catalytic NTP-binding sites ($\mu = 0.5$, $\mu_\sigma = 0.1$; $p <$

0.05; 2.7 Å; Figure 21b, Table S4a). (Analyses at 2.0 Å with solvent water gave similar results, Table S4b and Table S5b). Hydrogen bonds with the $O_{3\beta}$ most commonly involve neutral donors particularly from backbone nitrogens (Table S6). Other donors include water and positively charged amino acids, especially arginine. The donor is often within a conserved motif, for example at position 4 of the Walker A motif (p-loop; Table S6).

Interactions with $O_{3\alpha}$ in enzymes that cleave the P_α—$O_{3\alpha}$ bond are more frequent than in the comparison groups, but more marginally so. Only the difference between P_α—$O_{3\alpha}$ cleaving sites ($\mu = 0.4$, $\mu_\sigma = 0.1$) and $O_{3\beta}$—P_γ cleaving sites ($\mu = 0.3$, $\mu_\sigma < 0.1$) reach statistical significance ($p < 0.05$, Table S4a). With fewer enzymes exhibiting even a single interaction (Figure 21c, Table S4 and Table S5), the statistical criteria are harder to meet, and it is possible (but unproven) that $O_{3\alpha}$ interactions are as important as $O_{3\beta}$ interactions, but only for some subset of P_α—$O_{3\alpha}$ cleaving enzymes.

Experimental modulation of $O_{3\beta}$ interactions in Arginine Kinase – The results of the database analysis can be rationalized by hypothesizing a mechanistic role for interactions at the β-phosphoryl oxygen in $O_{3\beta}$—P_γ-cleaving enzymes. The hypothesis was tested in the arginine kinase (AK) model enzyme system. Phosphagen kinases have been a paradigm for study of the catalysis of multisubstrate reactions. AK was chosen in part due to the availability of a transition state analog crystal structure at 1.2 Å resolution[52] from which interactions of the $O_{3\beta}$ with Arg_{280} and Arg_{126} could be observed with precision. Arg_{280} also interacts with the α-nonbridging oxygen of the nucleotide substrate and an enzyme aspartate side chain, but not the γ-oxygens of the substrate (Figure 22). It was posited that interactions with $O_{3\beta}$ could be modulated with minimal

direct impact on the γ-phosphate. Of several mutants expressed and purified, the conservative substitution R280K proved most interesting (Table 2). Relative to the wild-type (WT), k_{cat} was decreased from 104 s^{-1} to $1s^{-1}$, whereas K_M values for the two substrates remained within 3-fold of WT (R280K: K_M(Arg) = 0.8 mM; K_M(ATP) = 0.5 mM. WT K_M(Arg) = 0.3 mM; K_M(ATP) = 0.3 mM).[125] The much larger impact upon rate indicates that the affect of the mutation is predominantly upon catalysis as opposed to substrate binding.

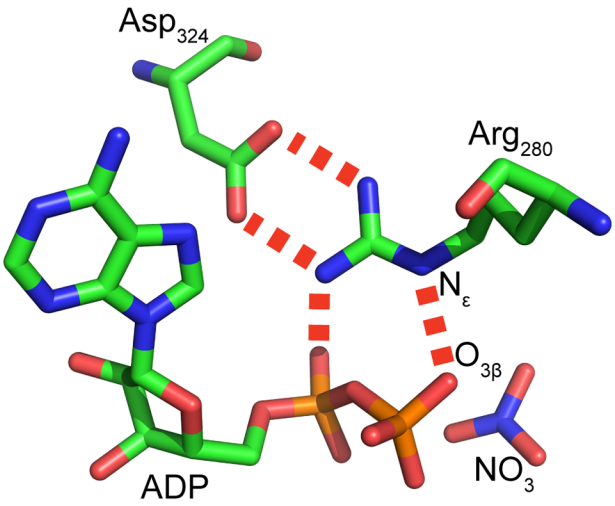


Figure 22. Arginine kinase active site mutation. Selected interactions of the nucleotide in the crystal structure of the transition state analog complex of Horseshoe Crab Arginine Kinase (AK, PDB ID 1M15[52]). Carbon = green, nitrogen = dark blue, oxygen = red and phosphorus = orange. In AK, Arg_{280} contacts the $O_{3\beta}$ oxygen of ADP, an α-oxygen and the Asp_{324} side chain. In ATP, the $O_{3\beta}$ oxygen bridges to the γ-phosphate which is mimicked by nitrate in this transition state analog complex. Hydrogen bonds are shown with red dotted lines.

Table 2. Kinetic analysis of the Arginine kinase R280K mutant.

K_s (ATP) mM	1.4 ± 0.2
K_M (ATP) mM	0.5 ± 0.3
K_s (Arg) mM	0.9 ± 0.1
K_M (Arg) mM	0.8 ± 0.3
α	0.7 ± 0.1
K_{cat} (s^{-1})	1.4 ± 0.1

Importance of $O_{3\beta}$ interactions reminds us of analogous interactions investigated specifically for Ras of the MAPK signal transduction cascade. Crystal structures and kinetic isotope studies of Ras alone and in complex with a GTPase activing protein (GAP), have highlighted interactions at the β-phosphoryl group of GTP as more important than those of the α- or γ-groups.[126,127] Upon cleavage of the γ-phosphate, electrons from the $O_{3\beta}$—P_γ bond are transferred to the $O_{3\beta}$ oxygen and distributed among the two non-bridging β-oxygens through resonance.[128] It was proposed that hydrogen bonds with $O_{3\beta}$ stabilize the accumulated charge: one from the arginine inserted into the active site by the GAP, and the other from a Ras backbone nitrogen.[128] The database analysis and kinetic characterization of AK mutants indicates that $O_{3\beta}$ interactions figure in a much broader and more diverse set of phosphoryl transferases than hitherto appreciated. The next section examines the potential catalytic impact of such interactions.

Quantum mechanical analysis of $O_{3\beta}$ interactions – In Ras-GAP, FTIR measurements show binding of GTP to Ras and introduction of the GAP, reduces β- and increases γ-phosphoryl group bond orders.[129] The bond order of the non-bridging P_β—O_β bonds is reduced further in going from ground to transition state, as determined by kinetic

isotope studies.[128] Implications are that the β-oxygens gain negative charge that originates from the γ-oxygens which are rendered more positive.[128,129] Also demonstrated is a decrease in $O_{3\beta}$—P_{γ} bond order suggesting a weakening/lengthening that may be important in the hydrolysis reaction.[128] It is thought that electrostatic stabilization of negative charge at the β-oxygens draws negative charge out of the $O_{3\beta}$—P_{γ} bond onto the $O_{3\beta}$ oxygen, which is then partially transferred to the non-bridging β-oxygens through resonance.[126] Presumably charge transfer would be achieved through induction. A small molecule quantum study of a series of organic phosphates indicated, however, that induction is unlikely to be a factor in scissile bond lengths, whereas hyperconjugation was found to be important.[15] Hyperconjugation is a stereoelectronic effect, where electron density is transferred from an electron rich donor orbital to the anti-bonding orbital of a neighboring bond.[72] An example is depicted in Figure 23, where electron density is transferred from a γ-oxygen lone pair orbital in methyl triphosphate into the antibonding orbital of the $O_{3\beta}$—P_{γ} bond.

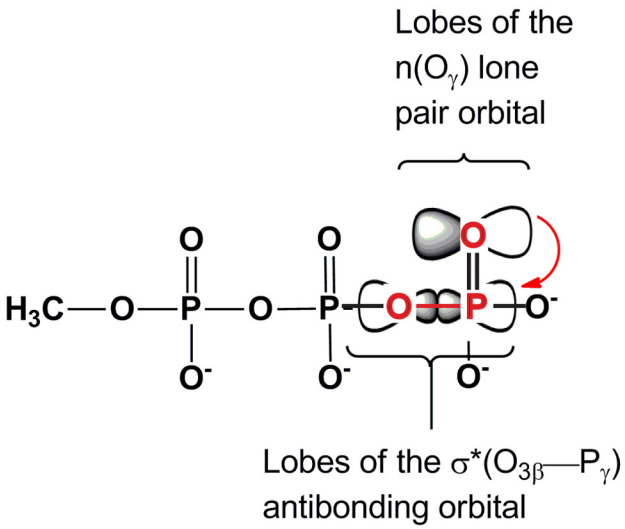


Figure 23. Hyperconjugation in methyl triphosphate, a model for ATP. Electron density is transferred from a lone pair orbital on one of the γ-oxygens (O_γ), into the antibonding orbital (σ) of the $O_{3\beta}$—P_γ bond. Red lettering indicates the atoms involved in the hyperconjugative interaction and the red arrow represents electron density transfer.*

The changes in charge and $O_{3\beta}$—P_γ bond length inferred through experiment are reminiscent of the effects of hyperconjugation on nucleotide chemistry seen in our past modeling studies.[15,124] Enhanced hyperconjugation would both render γ-oxygens more positive and $O_{3\beta}$ more negative as electron density is transferred from the γ-phosphoryl group to the $O_{3\beta}$—P_γ bond and would lengthen the scissile bond. The hypothesis that interactions with O_β enhance hyperconjugation in the nucleotide is tested here, using quantum mechanical calculations of model systems. (Figure 20). The relationship between the N—H•••$O_{3\beta}$ hydrogen bond length, hyperconjugation and the $O_{3\beta}$—P_γ bond length is investigated through perturbation of the N—H•••$O_{3\beta}$ hydrogen bond.

Density functional calculations on Structure 1 (Figure 20a) show that a single hydrogen bond, between N-methylacetamide and the $O_{3\beta}$ oxygen in methyl triphosphate, induces an increase in hyperconjugation and ~0.05 Å elongation of the $O_{3\beta}$—P_γ bond

(Figure 24). Our initial calculations indicate that it is at least plausible that the $O_{3\beta}$—P_{γ} bond length could be modulated by $O_{3\beta}$ hydrogen bonding by increasing hyperconjugation.

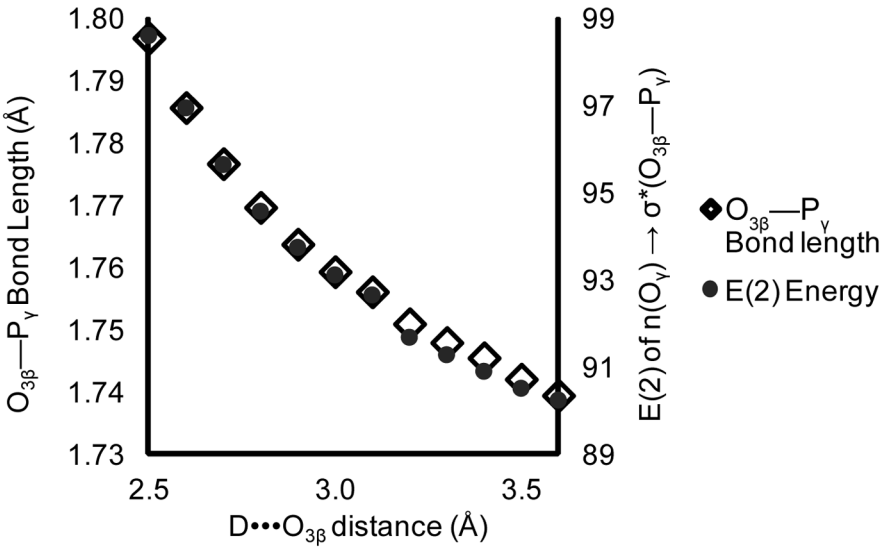


Figure 24. The $O_{3\beta}$—P_{γ} bond length and hyperconjugation increase with decreasing D•••$O_{3\beta}$ hydrogen bond length in Structure1. Calculated E(2) energy is of the n(O_{γ}) → σ($O_{3\beta}$—P_{γ}) hyperconjugative interaction. D is the N—H hydrogen bond donor in N-methylacetamide.*

Two factors determine the magnitude of hyperconjugation; the degree of donor and acceptor orbital overlap, and the orbital energy gap. We next dissect the impact of a hydrogen bond on these factors. Our previous study of organic phosphates revealed a positive correlation between hyperconjugation and polarity of the O—P bond in ten compounds[15].[15] This correlation resulted from a lowering of the σ*(O—P) antibonding orbital energy and therefore a decrease in the energy gap between the n(O) and σ*(O—P) orbitals (see Eq. 1). The orbital overlap (F_{ij} in Eq. 1) and the orbital energies of n(O_{γ}) and σ*($O_{3\beta}$—P_{γ}) were now similarly calculated for Structure 1 (Figure 20a), as a function of

hydrogen bond length (Figure 25). There is a strong dependence of the energy of the $\sigma^*(O_{3\beta}—P_\gamma)$ antibonding orbital upon strength of the hydrogen bond (Figure 25a). However, neither the energy of the non-bridging γ-oxygen lone pair orbitals (Figure 25b) nor the orbital overlap (Figure 25c) are affected. As the hydrogen bond is shortened (made stronger), the orbital energy gap decreases, increasing hyperconjugation and elongating the $O_{3\beta}—P_\gamma$ bond.

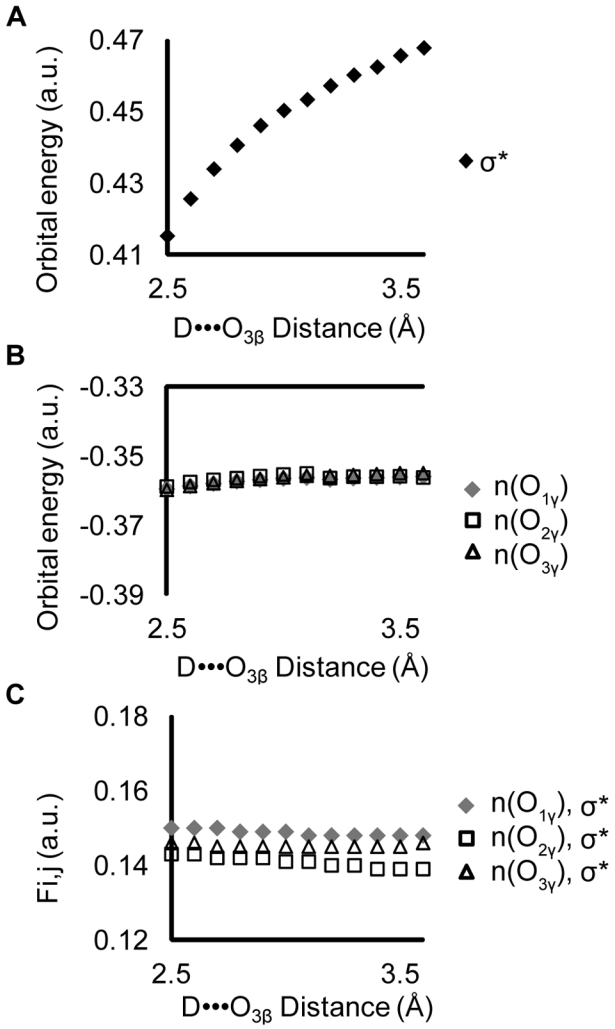


Figure 25. Effects of a hydrogen bond at $O_{3\beta}$ on orbital and interaction energies in Structure1. Shortening the hydrogen bond between N-methylacetamide and methyl triphosphate: a) decreases the orbital energy of the $\sigma^*(O_{3\beta}—P_\gamma)$ anti-bonding orbital; while (b) leaving unchanged the n(Oγ) donor orbital energies and (c) $F_{i,j}$, a measure of the overlap between the n(Oγ) lone pair orbitals and $\sigma^*(O_{3\beta}—P_\gamma)$. D denotes hydrogen bond donor. σ^* denotes $\sigma^*(O_{3\beta}—P_\gamma)$.

The relationship between $O_{3\beta}—P_\gamma$ bond elongation and hyperconjugation was confirmed by protonating the γ-oxygens. Protonation of the γ-oxygens has been shown to decrease the hyperconjugative interaction $n(O_\gamma) \rightarrow \sigma^*(O_{3\beta}—P_\gamma)$.[15,124] Protonation

decreases the dependence of the $O_{3\beta}$—P_γ bond length on the strength of the D•••$O_{3\beta}$ hydrogen bond (Figure 26). This is consistent with the impact of hydrogen bonding upon $O_{3\beta}$—P_γ elongation being mediated by hyperconjugation.

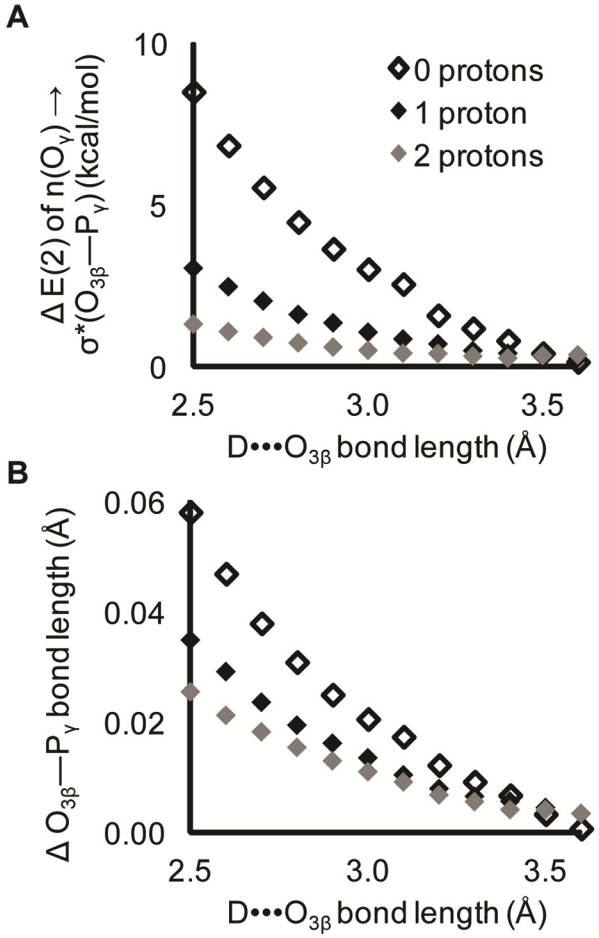


Figure 26. Impact of protonation of the γ-oxygens. Protons decrease the effects of the $O_{3\beta}$ hydrogen bond on both a) the magnitude of hyperconjugation and b) the $O_{3\beta}$—P_γ bond length. Changes are calculated relative to corresponding structures without a D•••$O_{3\beta}$ hydrogen bond. The ability of interactions with the $O_{3\beta}$ to increase the scissile bond length is reliant on a strong intrinsic hyperconjugation in the ligand. When hyperconjugation is discrupted through protonation, the ability is eliminated.

Hyperconjugation is also subject to "secondary" perturbations that might be of greater physiological relevance; hydrogen bonds added at the γ-oxygens interfere with hyperconjugation.[124] The primary dependence of $O_{3\beta}$—P_γ elongation on the strength of the $O_{3\beta}$ hydrogen bond was then recalculated, after the addition of hydrogen bonds at γ, to assess its dependence on the extent of secondary interactions. Hydrogen bond donors at the γ-oxygens from water (Figure 20d), acetamide (modeling an aspartate side chain, Figure 20e) and propylazanium (modeling a lysine, Figure 20f) were all examined. Hydrogen bonds to γ-oxygens (secondary interactions) were fixed at 2.8 Å, while the primary effect was modulated by varying the hydrogen bond between N-methylacetamide and the $O_{3\beta}$ oxygen from 2.5 Å to 3.1 Å in 0.2 Å increments.

A water or neutral acetamide hydrogen bond with a γ-oxygen has little effect (Figure 27a). Only the salt bridge bond from a positively charged propylazanium donor modulates the primary interactions. Even so, the effect is modest: a salt bridge between propylazanium and O_γ reduces the bond elongation by about 15%. It is possible that multiple salt bridges to γ-oxygens could have an additive effect. However, our database analysis shows (Figure 27b) fewer mean γ-oxygen salt bridges in $O_{3\beta}$—P_γ–cleavage sites (μ= 1.3, μ_σ = 0.1) than in non-catalytic sites (μ= 2.0, μ_σ = 0.2; t-test; $p < 0.05$; 2.7 Å structure set, detailed statistics in Table S7). In summary, analysis of the secondary interactions supports the notion that the effects of hydrogen bonding at $O_{3\beta}$ are modulated through hyperconjugation, and that enzymes have evolved environments around the γ-phosphate to limit interactions that reduce the hyperconjugation.

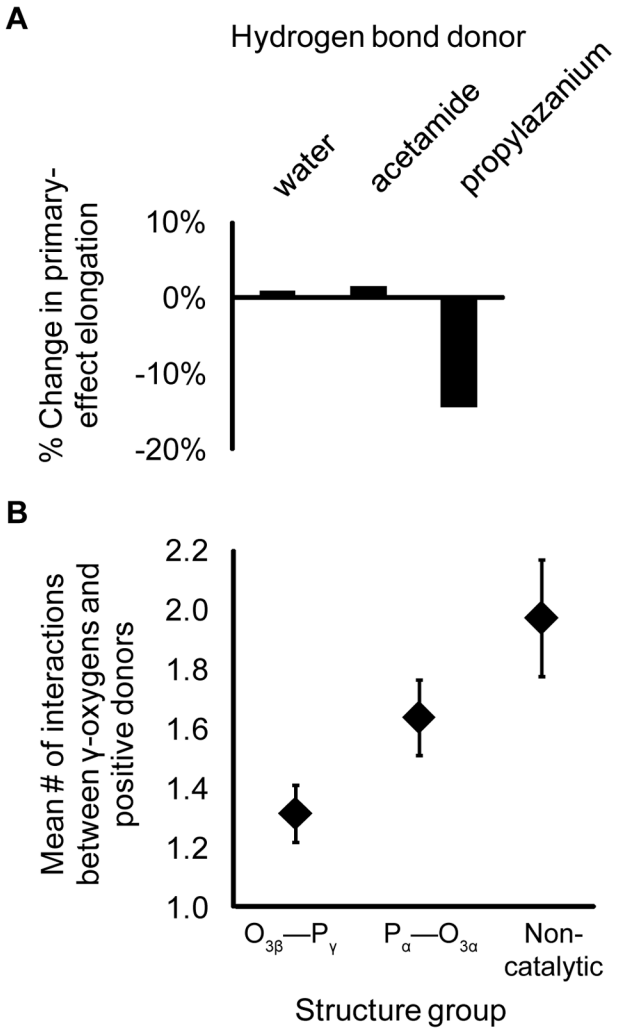


Figure 27. Impact of (secondary) interactions with γ-oxygens on $O_{3\beta}$—P_γ bond elongation induced by (primary) $O_{3\beta}$ interactions. a) Secondary interactions with γ-oxygens have modest impact and are greatest for charged donors. b) From the database analysis, the mean number of salt bridges with γ-oxygens are given for $O_{3\beta}$—P_γ-, P_α—$O_{3\alpha}$-cleaving active sites and non-catalytic sites. The resolution cutoff of structures used is 2.7 Å. Bars denote standard errors. There are fewer salt bridges at the γ-oxygens in enzymes that cleave $O_{3\beta}$—P_γ bonds relative to sites that bind an unreactive NTP.

Quantum mechanical analysis of interactions with non-bridging oxygens – Of interest is whether hyperconjugation can explain the higher frequencies of interactions

with non-bridging β-oxygens in $O_{3\beta}$—P_{γ} cleaving active sites, relative to non-catalytic binding sites (Figure 21). Similar to our study of $O_{3\beta}$ interactions, QM calculations were performed for N-methyl acetamide hydrogen-bonded to the β-nonbridging oxygen in Structure 7 (Figure 20g). Changes in the $n(O_{\gamma}) \rightarrow \sigma^*(O_{3\beta}—P_{\gamma})$ hyperconjugative interaction were calculated as the hydrogen bond length was varied from 2.5 Å to 3.5 Å in 0.2 Å increments.

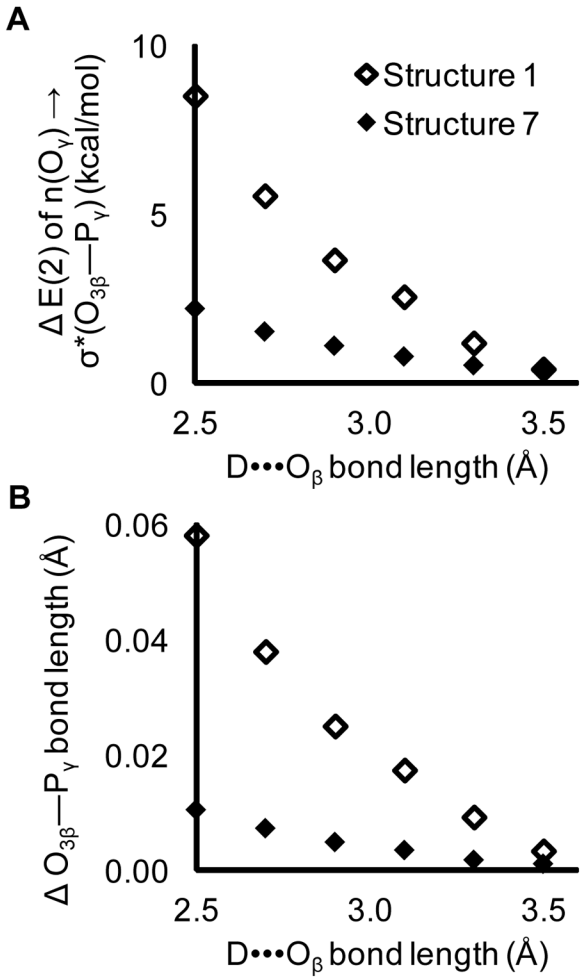


Figure 28. Impact of a hydrogen bond with a non-bridging β-oxygen. An interaction between N-methylacetamide and a non-bridging β-oxygen of methyl triphosphate decreases the (a) hyperconjugation and (b) scissile bond length. Interactions with the non-bridging β-oxygen (filled diamonds, Structure 7) are compared to those for $O_{3\beta}$ (open diamonds, Structure 1). Changes are measured relative to methyl triphosphate absent a hydrogen bond.

The impact of a single hydrogen bond with a non-bridging β-oxygen is smaller than that of an $O_{3\beta}$ hydrogen bond (Figure 28). However, the database analysis revealed an average of four enzyme interactions summed across the two non-bridging β-oxygens

(Figure 21a). Previous work has shown that multiple hydrogen bonds to the same oxygen in methyl triphosphate have an additive effect on hyperconjugation.[124] Thus the aggregate effect of interactions with non-bridging β-oxygens could be commensurate with those of the $O_{3\beta}$.

The catalytic impact of phosphate interactions has to be assessed relative to the uncatalyzed reaction in solution. The hydrogen bonding interactions described above that elongate the scissile bond are unlikely to exceed saturating interactions with water in free solution. How then, can the active site interactions be considered catalytic? In addressing this, it is noted that, whereas hydrogen bonding with the β-oxygens destabilizes the scissile $O_{3\beta}$—P_{γ} bond, interactions with the non-bridging γ-oxygens are known to stabilize the scissile bond through the same type of hyperconjugative effects.[124] Thus, the desolvation of the β- and γ-oxygens that occurs on substrate binding has opposite effects upon bond elongation. A net catalytic effect could be obtained (or at least a catalytic cost mitigated) through reduction in γ-oxygen interactions while maintaining a higher level of β-oxygen interaction. To test this prediction, the dataset of 2 Å structures was queried for the fractional saturation of lone pairs with enzyme or solvent interactions, comparing the total five lone pairs of non-bridging β-oxygens to the 8 of non-bridging γ-oxygens. The prediction is clearly borne out (Figure 29, detailed statistic given in Table S8), as is another prediction of a hyperconjugative mechanism, that in $O_{3\beta}$—P_{γ}-cleaving enzymes, there are more interactions with the bridging $O_{3\beta}$ than either the $O_{3\alpha}$ or $O_{5'}$. The favoring of interactions with the β-phosphoryl group that we see in the database analysis is supported by the earlier FTIR studies of Ras that show

stronger binding of the β-phosphoryl group.[130] Together the results make a strong case for an extensive role of hyperconjugation in enzyme-catalyzed phosphoryl transfer.

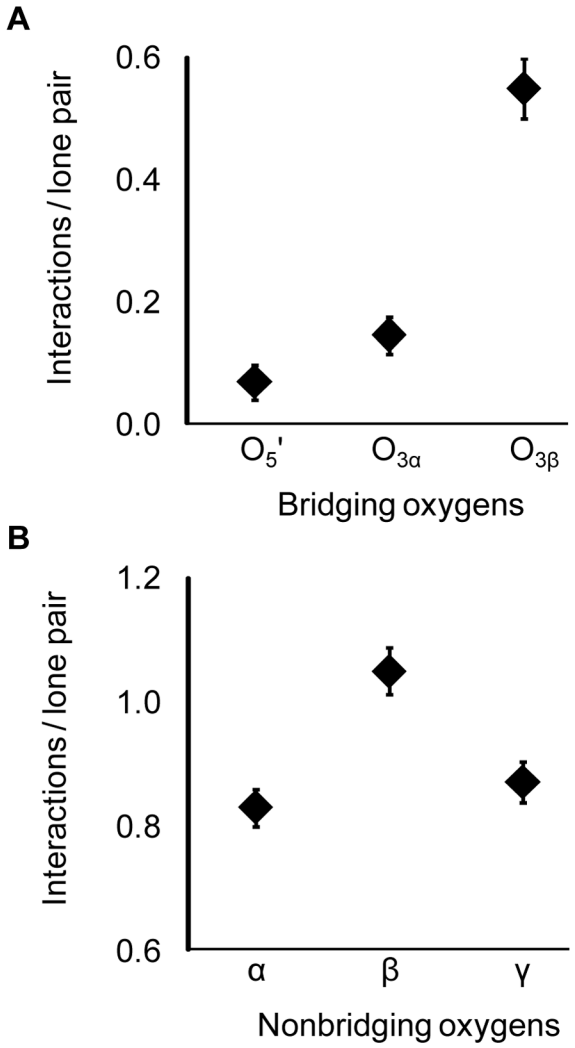


Figure 29. Saturation of NTP oxygen lone pairs by enzyme or solvent interactions in $O_{3\beta}$—P_{γ}-cleaving active sites. Bars denote standard errors.

DISCUSSION

Phosphoryl transfer occurs in a diverse set of functional contexts, including metabolism, transport, motility and signal transduction. Classical mechanistic analyses have revealed the details of selected systems, highlighting potential differences, but identifying few common themes among the diverse enzyme families. This work

represents a case study of a complementary top-down analysis, testing whether the statistics of large samples of structures can reveal unifying principles, even in the presence of experimental error in individual structures and likely real differences between the mechanisms of individual enzymes.

The prevalence of some binding site interactions was dependent on which, if any, of the phosphoanhydride bonds was cleaved in a reaction. One of these interactions, a hydrogen bond to the $O_{3\beta}$, had been observed in Ras.[131] Its pervasiveness in the database analysis indicates that this type of interaction is important not only to Ras and relatives, but a broad spectrum of phosphoryl transferases. Support both for this enzymological inference (and for the approach of top-down analysis) comes from the isolation of a partially active mutant of an unrelated kinase, in which conservative substitution at the site of interaction decreases experimentally measured rate with little impact upon binding.

Ground state quantum mechanical modeling provides a unifying rationalization for such $O_{3\beta}$ interactions. In support are the changes in bond order and partial charge inferred from kinetic isotope and FTIR studies,[128,129] together with additional preferences in β- over the γ-oxygen interactions that can be resolved in the database analysis. The common denominator appears to be hyperconjugative interactions involving the σ* antibonding orbital of the scissile bond. Hydrogen bonding at the bridging $O_{3\beta}$ and non-bridging β-oxygens closes the energy orbital gap, enhancing the hyperconjugative elongation of the $O_{3\beta}$—P_{γ} bond. Interactions at the more distal γ-oxygens have the opposite impact upon hyperconjugation, explaining why fewer interactions are observed with the γ-phosphate.

The current analysis examines only the ground and not the transition state, and does not attempt to model the chemical reaction either in the enzyme or free solution, studies that would be required for definitive mechanistic conclusions. However, the prevalence of similar interactions visible in the structures of diverse phosphoryl transferases suggests that there might have been evolutionary selective pressure to modulate underlying electronic interactions within the substrates in fundamentally consistent ways. Furthermore, the current analysis cannot rank the importance of the different mechanistic strategies that might be combined in an individual enzyme. The current approach will highlight interactions that are common among diverse enzymes, which may or may not be the most critical for an individual enzyme. For arginine kinase, the 100-fold change in rate upon conservative mutational substitution for Arg_{280} is of the same order as that observed on mutation of the catalytic base that interacts with the non-nucleotide substrate [7] and 50-fold less than mutations to a cysteine thiolate that stabilizes the electrostatics of the guanidinium in the presumptive transition state [8]. Thus, the work here paints a picture of phosphagen kinases employing several means of enhancing rate, of which the hyperconjugative interactions investigated here are an important, but not dominant component. For other enzymes, the relative importance of hyperconjugation among other catalytic mechanisms is not known quantitatively. Nevertheless, the pervasiveness of hydrogen-bonding motifs that would favorably modulate catalysis is indicative of some importance to a wide range of enzymes, making such interactions worthy of additional study.

ACKNOWLEDGMENTS

Omar Davulcu is thanked for help in protein purification and enzyme kinetics and Nancy Meyer is thanked for help with generating figures. The authors would also like to thank Johannes Elferich and Kyle Ambert for help with R and Python scripting.

SUPPLEMENTAL DATA

Table S1. Protein Data Bank structures used in study. As some structures contain multiple ligands, the specific ligand used in this study is specified.

PDB ID	Structure Title	Source	Ligand Name	Ligand Chain	Ligand Number
	Proteins with non-catalytic binding sites				
1BCP	Pertussis toxin	Bordetella pertussis	ATP	E	111
1C4K	Ornithine decarboxylase	Lactobacillus sp. 30A	GTP	A	999
1GZ3	NAD-dependent malic enzyme	Homo sapiens	ATP	A	601
1H78	Anaerobic ribonucleotide-triphosphate reductase large chain	Bacteriophage T4	DCP	A	1588
1M7B	Rnd3/RhoE small GTP-binding protein	Homo sapiens	GTP	A	538
1RA7	Poliovirus Polymerase with GTP	Human poliovirus 1	GTP	A	1001
1UDW	Uridine-cytidine kinase 2	Human	CTP	A	301
1V3S	Nitrogen regulatory protein P-II	Thermus thermophilus	ATP	A	200
1WUR	GTP cyclohydrolase I	Thermus thermophilus	8DG	A	821
1XBT	Thymidine kinase, cytosolic	Homo sapiens	TTP	A	1195
1XJK	ribonucleotide reductase, B12-dependent	Thermotoga maritima	DGT	A	1003
4FYW	E. coli Adenylate cyclase	Escherichia coli	CTP	B	202
4FYY	Aspartate carbamoyltransferase catalytic chain	Escherichia coli	UTP	B	202
2BE9	Aspartate carbamoyltransferase	archaeon Sulfolobus acidocaldarius	CTP	B	401
2BNF	Uridylate kinase	Escherichia coli	UTP	A	1242
2CVW	Ribonucleoside-diphosphate reductase large chain 1	Saccharomyces cerevisiae	TTP	A	1001
2E5Y	ATP synthase epsilon chain	Bacillus sp. PS3	ATP	A	1001
2EG2	Nitrogen regulatory protein P-II	Aquifex aeolicus	ATP	A	501
2GNK	GLNK, nitrogen regulatory protein	Escherichia coli	ATP	A	200
2HS0	Phosphoribosylformylglycinamidine synthase II	Thermotoga maritima	ATP	A	1065
2HVW	deoxycytidylate deaminase	Streptococcus mutans	DCP	A	1201
2ILZ	poliovirus polymerase	Human poliovirus 1	GTP	A	1001
2J7W	Dengue virus NS5 RNA dependent RNA polymerase domain	Dengue virus	GTP	A	1886
2J9C	GLNK1, nitrogen regulatory PII-like protein	Methanococcus jannaschii	ATP	A	1116
2J9L	Chloride channel protein 5	Homo sapiens	ATP	A	1752
2PNN	Transient receptor potential cation channel subfamily V member 1	Rattus norvegicus	ATP	A	374
2QXX	Deoxycytidine triphosphate deaminase	Mycobacterium tuberculosis	TTP	A	201
2RD5	Acetylglutamate kinase-like protein	Arabidopsis thaliana	ATP	C	1000
2UZ3	Thymidine kinase	Ureaplasma urealyticum	TTP	A	1220
2V9J	5'-AMP-activated protein kinase catalytic subunit alpha-1	Rattus norvegicus	ATP	E	1328
2VP0	Deoxynucleoside kinase	Drosophila melanogaster	TTP	A	1210
2WGH	Ribonucleoside-diphosphate reductase large subunit	Homo sapiens	DTP	A	1745
2XJE	Cytosolic purine 5'-nucleotidase	Homo sapiens	ATP	A	1490
2XZW	Nitrogen regulatory protein P-II	Synechococcus elongatus	ATP	A	200
2YCH	Competence protein PILM	Thermus thermophilus	ATP	A	502
3CQD	6-phosphofructokinase isozyme 2	Escherichia coli	ATP	A	313
3EK5	Uridylate kinase	Xanthomonas campestris pv. campestris	GTP	E	2001
3HNC	Ribonucleoside-diphosphate reductase large subunit	Homo sapiens	TTP	A	802
3IRH	HD domain protein	Enterococcus faecalis	DGT	A	458
3LF0	Nitrogen regulatory protein P-II	Mycobacterium tuberculosis	ATP	A	701
3LFZ	Protein MJ1225 , a putative archaeal homolog of g-AMPK.	Methanocaldococcus jannaschii	ATP	A	281
3MHY	PII-like protein Pz	Azospirillum brasilense	ATP	A	113
3MW9	Glutamate dehydrogenase 1	Bos taurus	GTP	A	503
3NCQ	Nitrogen regulatory protein P-II (GlnB-2)	Archaeoglobus fulgidus	ATP	A	201
3NWY	Uridylate kinase	Mycobacterium tuberculosis	GTP	A	262
3O0O	Ribonucleoside-diphosphate reductase	Thermotoga maritima	TTP	A	1001
3TA2	Nitrogen regulatory protein P-II (GlnB-3)	Archaeoglobus fulgidus	ATP	A	121
3TUV	Insulin-degrading enzyme	Rattus norvegicus	ATP	A	1020
4A7W	Uridylate kinase	Helicobacter pylori	GTP	A	1241
1QHA	Hexokinase type I	Homo sapiens	ANP	A	999
	P_α—$O_{3\alpha}$–cleaving enzymes				
1EE1	NH(3)-dependent NAD(+) synthetase	Bacillus subtilis	ATP	A	5000
1F9A	NMN adenylyltransferase	Methanocaldococcus jannaschii	ATP	A	700
1FA0	Poly(A)-polymerase	Saccharomyces cerevisiae	3AT	A	604
1GN8	Phosphopantetheine adenylyltransferase	Escherichia coli	ATP	A	700
1GQ9	3-Deoxy-manno-octulosonate cytidylyltransferase	Escherichia coli	CTP	A	1243

1GX5	RNA-directed RNA polymerase	Hepatitis C virus	GTP	A	1531
1H7G	3-deoxy-manno-octulosonate cytidyltltransferase	Escherichia coli	CTP	B	1243
1I52	4-diphosphocytidyl-2-C-methylerythritol synthase	Escherichia coli	CTP	A	1001
1KNY	Kanamycin nucleotidyltransferase	Staphylococcus aureus	APC	A	556
1LV5	DNA polymerase I	Geobacillus stearothermophilus	DCP	A	201
1MAU	tryptophan-tRNA ligase	Geobacillus stearothermophilus	ATP	A	400
1MB9	Beta-lactam synthetase	Streptomyces clavuligerus	ATP	B	702
1MC3	Glucose-1-phosphate thymidylyltransferase	Escherichia coli	TTP	A	294
1N2B	Pantothenate synthetase	Mycobacterium tuberculosis	APC	A	801
1N35	Minor core protein lambda 3	Mammalian orthoreovirus 3	CH1	A	1291
1N77	Glutamyl-tRNA synthetase	Thermus thermophilus	ATP	A	600
1NB6	RNA polymerase (HC-J4)	Hepatitis C virus	UTP	A	1700
1Q19	Carbapenam Synthetase	Pectobacterium carotovorum	APC	A	505
1Q79	Poly(A) polymerase alpha	Bos taurus	3AT	A	1000
1RN8	Deoxyuridine 5'-triphosphate nucleotidohydrolase	Escherichia coli	DUP	A	777
1T7P	T7 DNA polymerase	Enterobacteria phage T7	DG3	P	23
1V25	long-chain-fatty-acid-CoA synthetase	Thermus thermophilus	ANP	A	666
1WC5	Adenylate cyclase	Spirulina platensis	APC	A	1500
1WVC	Glucose-1-phosphate cytidylyltransferase	Salmonella enterica subsp. enterica serovar Typhi	CTP	A	401
1XDN	RNA editing ligase MP52	Trypanosoma brucei	ATP	A	501
1XNG	NH(3)-dependent NAD(+) synthetase	Helicobacter pylori	ATP	A	303
1YBU	adenylyl cyclase Rv1900c CHD	Mycobacterium tuberculosis	APC	A	601
1YFR	Alanyl-tRNA synthetase	Aquifex aeolicus	ATP	A	1500
1YID	tryptophanyl-tRNA synthetase	Deinococcus radiodurans	ATP	B	3000
1ZUN	Sulfate adenylyltransferase subunit 2	Pseudomonas syringae	AGS	A	600
2AGQ	DNA polymerase IV	Sulfolobus solfataricus	DTP	A	1814
2AQ4	DNA repair protein REV1	Saccharomyces cerevisiae	DCP	A	201
2ARU	Lipoate-protein ligase A	Thermoplasma acidophilum	ATP	A	1065
2B56	RNA editing complex protein MP57	Trypanosoma brucei	UTP	A	501
2BT1	Deoxyuridine 5'-triphosphate nucleotidohydrolase	Human herpesvirus 4	DUP	A	1257
2BZ0	GTP cyclohydrolase II	Escherichia coli	G2P	A	1176
2C5U	T4 RNA Ligase (RNL1)	Bacteriophage T4	APC	A	1375
2CIC	Deoxyuridine 5'-triphosphate nucleotide hydrolase	Campylobacter jejuni	DUP	A	1968
2CJA	Seryl-tRNA synthetase	Methanosarcina barkeri	ATP	A	1505
2FAQ	LigD polymerase	Pseudomonas aeruginosa	ATP	A	1304
2FMP	DNA polymerase beta	Homo sapiens	DCT	A	1338
2GA9	Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase	Vaccinia virus	AGS	D	482
2HQU	Deoxyuridine 5'-triphosphate nucleotidohydrolase	Homo sapiens	DUP	A	777
2I4O	Proline-tRNA ligase	Rhodopseudomonas palustris	ATP	A	442
2IHM	DNA polymerase mu	Mus musculus	D3T	A	638
2IKF	RNA uridylyl transferase	Trypanosoma brucei	UTP	A	501
2IRY	DNA ligase-like protein Rv0938/MT0965	Mycobacterium tuberculosis	DGT	A	1102
2J3M	Prolyl-tRNA synthetase	Enterococcus faecalis	ATP	A	701
2OKE	Deoxyuridine 5'-triphosphate nucleotidohydrolase	Vaccinia virus	DUP	B	201
2PFO	DNA polymerase lambda	Homo sapiens	DUP	A	952
2PI4	DNA-directed RNA polymerase	Enterobacteria phage T7	GH3	A	902
2PY4	Deoxyuridine 5'-triphosphate nucleotidohydrolase	Mycobacterium tuberculosis	DUP	A	777
2PYL	Phi29 DNA polymerase	Bacillus phage phi29	TTP	A	3204
2Q0F	RNA uridylyl transferase	Trypanosoma brucei	UTP	A	501
2QUI	Tryptophanyl-tRNA synthetase	Homo sapiens	ATP	A	502
2W02	Achromobactin synthetase protein D (ACSD)	Erwinia chrysanthemi	ATP	A	1588
2XCE	Probable deoxyuridine 5'-triphosphate nucleotidylhydrolase YNCF	Bacillus subtilis	DUP	C	1131
2XI3	RNA-directed RNA polymerase	Hepatitis C virus	GTP	A	1001
2XTI	Asparaginyl-tRNA synthetase	Brugia malayi	ATP	B	1549
2XWL	2-C-Methyl-D-erythritol 4-phosphate cytidylyltransferase	Mycobacterium smegmatis	CTP	A	301
2XY3	Prophage DUTPase	Bacillus subtilis	DUP	C	1129
2Y27	Phenylacetate-coenzyme A ligase	Burkholderia cenocepacia	ATP	A	1440
2Y6P	3-Deoxy-manno-octulosonate cytidylyltransferase	Aquifex aeolicus	CTP	A	1233
2YAY	DUTPASE	Leishmania major	DUP	A	1265
2ZCE	Pyrrolysyl-tRNA synthetase	Methanosarcina mazei	ANP	A	501
2ZUE	Arginyl-tRNA synthetase	Pyrococcus horikoshii	ANP	A	700
3BJU	Lysyl-tRNA synthetase	Homo sapiens	ATP	A	603

3BQ1	DNA polymerase IV	Sulfolobus acidocaldarius	DG3	P	11
3BSO	RNA dependent RNA polymerase	Norwalk virus	CTP	A	514
3F2B	DNA-directed DNA polymerase III alpha chain	Geobacillus kaustophilus	DGT	A	1456
3FCE	D-alanine--poly(phosphoribitol) ligase subunit 1	Bacillus cereus	ATP	A	711
3G5A	FMN adenylyltransferase	Candida glabrata	APC	C	305
3GQC	DNA repair protein REV1	Homo sapiens	DCP	A	101
3H4D	DNA polymerase iota	Homo sapiens	DGT	T	875
3H5Y	RNA dependent RNA polymerase	Norwalk virus	CTP	A	515
3HIY	Minor Editosome-Associated TUTase	Trypanosoma brucei	UTP	A	501
3HXY	Alanyl-tRNA synthetase	Escherichia coli	ACP	A	444
3IAY	DNA polymerase delta catalytic subunit	Saccharomyces cerevisiae	DCP	A	986
3JZ0	Lincosamide nucleotidyltransferase	Enterococcus faecium	APC	A	738
3K59	DNA polymerase II	Escherichia coli	DCP	A	914
3K8D	3-deoxy-manno-octulosonate cytidylyltransferase	Escherichia coli	CTP	A	1243
3KTQ	DNA POLYMERASE I	Thermus aquaticus	DCT	A	113
3LWL	Klenow fragment of Taq polymerase	Thermus aquaticus	DDS	A	1
3MEY	class II aaRS homologue (Bll0957)	Bradyrhizobium japonicum	ATP	A	1010
3MFH	DNA polymerase eta	Saccharomyces cerevisiae	DTP	A	514
3N0Y	Adenylate cyclase 2	Yersinia pestis	APC	A	181
3NBA	Phosphopantetheine adenylyltransferase	Mycobacterium tuberculosis	APC	A	200
3NCI	RB69 DNA Polymerase	Enterobacteria phage RB69	DCP	A	904
3NEM	Aspartyl-tRNA synthetase	Thermococcus kodakarensis	ATP	B	1039
3OVB	CCA-Adding Enzyme	Archaeoglobus fulgidus	ATP	A	501
3Q23	Virion RNA polymerase	Enterobacteria phage N4	G2P	A	1108
3REU	AsnS-like asparaginyl-tRNA synthetase related protein	Pyrococcus abyssi	ATP	A	298
3TI0	DNA polymerase I	Geobacillus sp.	DG3	A	201
3TSU	Transcriptional regulatory protein	Escherichia coli	ANP	A	748
3TUX	RNA 3'-terminal phosphate cyclase	Escherichia coli	ATP	A	501
3VES	O-carbamoyltransferase TobZ	Streptoalloteichus tenebrarius	APC	A	602
4DQI	DNA polymerase I	Geobacillus kaustophilus	DCP	D	902
4ECR	DNA polymerase eta	Homo sapiens	DTP	P	504
4EDK	DNA primase	Staphylococcus aureus	GTP	A	504
4FH5	Poly(A) RNA polymerase protein cid1	Schizosaccharomyces pombe	UTP	A	401
	$O_{3\beta}$—P_{γ}-cleaving enzymes				
1A5U	Pyruvate kinase	Oryctolagus cuniculus	ATP	A	535
1A82	Dethiobiotin synthetase	Escherichia coli	ATP	A	802
1AD5	Haematopoetic cell kinase HCK	Homo sapiens	ANP	A	1
1AM4	P50-RHOGAP	Homo sapiens	GNP	D	678
1AQ2	Phosphoenolpyruvate carboxykinase	Escherichia coli	ATP	A	541
1B23	Elongation factor TU	Escherichia coli + Thermus aquaticus	GNP	P	406
1B63	MUTL	Escherichia coli K12	ANP	A	380
1BIF	6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase	Rattus norvegicus	AGS	A	500
1BXR	Carbamoyl-phosphate synthase	Escherichia coli	ANP	A	1083
1C1Y	Ras-related protein RAP-1A	Homo sapiens	GTP	A	170
1CDK	Camp-dependent protein kinase	Sus scrofa	ANP	A	400
1CM8	Phosphorylated MAP kinase P38-gamma	Homo sapiens	ANP	A	400
1CSN	Casein kinase-1	Schizosaccharomyces pombe	ATP	A	299
1D2N	N-ethylmaleimide-sensitive fusion protein	Cricetulus griseus	ANP	A	1
1DAH	Dethiobiotin synthetase	Escherichia coli	ACP	A	226
1E4G	Cell division protein FTSA	Thermotoga maritima	ATP	T	500
1ESN	Pantothenate kinase	Escherichia coli	ANP	A	401
1EZ1	Phosphoribosylglycinamide formyltransferase 2	Escherichia coli	ANP	A	400
1F2U	RAD50 ABC-ATPase	Pyrococcus furiosus	ATP	A	901
1G17	Ras-related protein SEC4	Saccharomyces cerevisiae	GNP	A	201
1G3R	Cell division inhibitor	Pyrococcus furiosus	ACP	A	238
1G7T	Translation initiation factor IF2/EIF5B	Methanothermobacter thermautotrophicus	GNP	A	601
1GJV	[3-methyl-2-oxobutanoate dehydrogenase [lipoamide]] kinase	Rattus norvegicus	SAP	A	1383
1GS5	Acetylglutamate kinase	Escherichia coli	ANP	A	1260
1H74	Homoserine kinase	Methanocaldococcus jannaschii	SAP	B	400
1H7U	Mismatch repair endonuclease PMS2	Homo sapiens	SAP	A	1366
1I58	Chemotaxis protein CHEA	Thermotoga maritima	ACP	A	998
1I6I	Kinesin-like protein KIF1A	Mus musculus	ACP	A	500
1ID0	PHOQ histidine kinase	Escherichia coli	ANP	A	487

1II0	Arsenical pump-driving ATPase	Escherichia coli	ATP	A	591
1IR3	Insulin receptor	Homo sapiens	ANP	A	300
1J7U	Aminoglycoside 3'-phosphotransferase	Enterococcus faecalis	ANP	A	303
1JST	Cyclin-dependent kinase-2	Homo sapiens	ATP	A	300
1K5D	Ran-GPPNHP-RanBP1-RanGAP complex	Homo sapiens	GNP	A	1250
1KHB	Phosphoenolpyruvate carboxykinase, cytosolic (GTP)	Homo sapiens	GCP	A	704
1KO5	Gluconate kinase	Escherichia coli	ATP	A	302
1KVK	mevalonate kinase	Rattus norvegicus	ATP	A	535
1LII	adenosine kinase	Toxoplasma gondii	ACP	A	799
1LP4	Protein kinase CK2	Zea mays	ANP	A	340
1M2O	protein transport protein SEC23	Saccharomyces cerevisiae	GNP	B	5200
1MH1	RAC1	Homo sapiens	GNP	A	200
1MX0	Type II DNA topoisomerase VI subunit B	Sulfolobus shibatae	ANP	A	900
1N6H	Ras-related protein Rab-5A	Homo sapiens	GNP	A	200
1NGJ	Heat-shock cognate 70 kD protein	Bos taurus	ANP	A	486
1NHI	DNA mismatch repair protein mutL	Escherichia coli K12	ANP	A	380
1NN5	thymidylate kinase	Homo sapiens	ANP	A	303
1O6L	Rac-beta serine/threonine protein kinase	HOMO SAPIENS	ANP	A	1480
1O6Y	Catalytic domain of PKNB kinase	Mycobacterium tuberculosis	ACP	A	1279
1OXV	GlcV, the ABC-ATPase of the glucose ABC transporter	Sulfolobus solfataricus	ANP	A	1104
1OYY	ATP-dependent DNA helicase	Escherichia coli	AGS	A	527
1P3D	UDP-N-acetylmuramate--alanine ligase	Haemophilus influenzae	ANP	A	603
1PVG	DNA topoisomerase II	Saccharomyces cerevisiae	ANP	A	901
1PYX	Glycogen synthase kinase-3 beta	Homo sapiens	ANP	A	1001
1Q97	SR protein kinase	Saccharomyces cerevisiae	ATP	A	485
1R2Q	Ras-related protein Rab-5A	Homo sapiens	GNP	A	200
1S16	Topoisomerase IV subunit B	Escherichia coli	ANP	A	1500
1SVM	SV40 large T antigen helicase domain	Simian virus 40	ATP	A	800
1SX3	groEL protein	Escherichia coli	ATP	A	1
1TID	Anti-sigma F factor	Geobacillus stearothermophilus	ATP	A	200
1TZ6	aminoimidazole riboside kinase	Salmonella typhimurium LT2	ACP	A	401
1W2C	Inositol-trisphosphate 3-kinase A	Homo sapiens	ANP	A	1462
1W5A	Cell division protein FTSZ homolog 1	Methanocaldococcus jannaschii	GTP	A	500
1W5T	ORC2	Aeropyrum pernix	ANP	A	700
1WUU	Galactokinase	Homo sapiens	ANP	C	395
1XDP	Polyphosphate kinase	Escherichia coli	ATP	A	701
1XR1	Proto-oncogene serine/threonine-protein kinase Pim-1	Homo sapiens	ANP	A	501
1Y8P	[Pyruvate dehydrogenase [lipoamide]] kinase isozyme 3	Homo sapiens	ATP	A	504
1YTM	phosphoenolpyruvate carboxykinase [ATP]	Anaerobiospirillum succiniciproducens	ATP	A	541
1Z2P	inositol 1,3,4-trisphosphate 5/6-kinase	Entamoeba histolytica	ACP	X	1476
1ZAO	Rio2 serine kinase	Archaeoglobus fulgidus	ATP	A	286
1ZP9	Rio1 kinase	Archaeoglobus fulgidus	ATP	A	260
1ZXM	DNA topoisomerase II, alpha isozyme	Homo sapiens	ANP	A	901
2A19	Eukaryotic translation initiation factor 2 alpha subunit	Saccharomyces cerevisiae	ANP	B	1640
2A2D	N-acetylgalactosamine kinase	Homo sapiens	ANP	A	461
2AFK	Nitrogenase molybdenum-iron protein	Azotobacter vinelandii	ACP	E	1292
2AJ4	Galactokinase	Saccharomyces cerevisiae	ANP	A	532
2BMU	Uridylate kinase	Pyrococcus furiosus	ANP	A	1226
2C49	Nucleoside kinase	Methanococcus jannaschii	ANP	A	1304
2DB3	ATP-dependent RNA helicase vasa	Drosophila melanogaster	ANP	A	2901
2DDO	Pyridoxine kinase	Escherichia coli	ATP	A	285
2DWP	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	Homo sapiens	ACP	A	700
2E8A	Heat shock 70kDa protein 1A	Homo sapiens	ANP	A	601
2EWS	Pantothenate kinase	Staphylococcus aureus subsp. aureus	ANP	A	1001
2FPM	DNA repair and recombination protein radA	Methanococcus voltae	ANP	A	401
2GJK	Upf1 helicase core	Homo sapiens	ANP	A	990
2GS7	EGFR kinase	Homo sapiens	ANP	A	301
2HF9	Probable hydrogenase nickel incorporation protein hypB	Methanocaldococcus jannaschii	GSP	A	300
2HS4	Phosphoribosylformylglycinamidine synthase II	Thermotoga maritima	ACP	A	900
2IO7	Bifunctional glutathionylspermidine synthetase/amidase	Escherichia coli	ANP	A	964
2IS4	DNA helicase II	Escherichia coli	ANP	A	700
2J0L	Focal adhesion kisase 1	Gallus gallus	ANP	A	1689
2J4J	Uridylate kinase	Sulfolobus solfataricus	ACP	D	228
2JG1	Tagatose-6-phosphate kinase	Staphylococcus aureus	ANP	A	1312

2JLV	Serine protease subunit NS3	Dengue virus 4	ANP	A	1619
2P0C	Proto-oncogene tyrosine-protein kinase MER	Homo sapiens	ANP	A	1
2PUN	Methylthioribose kinase	Bacillus subtilis	ACP	A	999
2PZE	Cystic fibrosis transmembrane conductance regulator	Homo sapiens	ATP	A	1
2Q7D	Inositol-tetrakisphosphate 1-kinase	Homo sapiens	ANP	A	917
2QEY	Phosphoenolpyruvate carboxykinase, cytosolic [GTP]	Rattus norvegicus	GTP	A	2696
2QF7	Pyruvate carboxylase protein	Rhizobium etli	AGS	A	1162
2QO9	Ephrin receptor	Homo sapiens	ANP	A	948
2QQ0	Thymidine kinase	Thermotoga maritima	ANP	A	601
2QT0	riboside kinase 1	Homo sapiens	ANP	A	1102
2QTV	Protein transport protein SEC23	Saccharomyces cerevisiae	GNP	B	200
2VOR	Folylpolyglutamate synthase protein FOLC	Mycobacterium tuberculosis	ACP	A	1490
2VPQ	Acetyl-COA carboxylase	Staphylococcus aureus	ANP	A	1449
2W00	HSDR subunit of the ECOR124I restriction enzyme	Escherichia coli	ATP	A	1886
2WJI	Ferrous iron transport protein B homolog	Methanocaldococcus jannaschii	GNP	A	1166
2X9H	Myosin-2 heavy chain	Dictyostelium discoideum	AD9	A	1001
2XAN	Inositol-pentakisphosphate 2-kinase	Arabidopsis thaliana	ANP	A	600
2Z1U	Hydrogenase expression/formation protein HypE	Desulfovibrio vulgaris subsp. vulgaris	ATP	A	342
3A1C	P- and N-domains of CopA, a copper-transporting P-type ATPase	Archaeoglobus fulgidus	ACP	A	997
3A4L	L-seryl-tRNA(Sec) kinase	Methanocaldococcus jannaschii	ANP	A	301
3AR2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	Oryctolagus cuniculus	ACP	A	1002
3ATT	Rv3168, a putative aminoglycoside antibiotics resistance enzyme	Mycobacterium tuberculosis	ATP	A	510
3C1M	threonine-sensitive aspartokinase	Methanocaldococcus jannaschii	ANP	A	472
3C9T	Thiamine monophosphate kinase	Aquifex aeolicus	ACP	A	307
3CQD	6-phosphofructokinase isozyme 2	Escherichia coli	ATP	B	312
3DCB	Kinesin-like protein Nod	Drosophila melanogaster	ANP	A	401
3DT7	Phosphoenolpyruvate carboxykinase, cytosolic [GTP]	Rattus norvegicus	GTP	A	800
3E7E	Mitotic checkpoint serine/threonine-protein kinase BUB1	Homo sapiens	ATP	A	1501
3EHG	Sensor kinase (YocF protein)	Bacillus subtilis	ATP	A	1
3EQD	Dual specificity mitogen-activated protein kinase kinase 1	Homo sapiens	AGS	A	2
3ETH	Phosphoribosylaminoimidazole carboxylase ATPase subunit	Escherichia coli	ATP	A	400
3EW9	DNA repair and recombination protein radA	Methanococcus maripaludis	ANP	A	401
3F5M	6-phospho-1-fructokinase (ATP-dependent phosphofructokinase)	Trypanosoma brucei	ATP	A	1001
3F5U	Death-associated protein kinase 1	Homo sapiens	ANP	A	296
3FHT	ATP-dependent RNA helicase DDX19B	Homo sapiens	ANP	A	480
3G3R	Vacuolar transporter chaperone 4	Saccharomyces cerevisiae	ANP	A	1000
3GON	Phosphomevalonate kinase	Streptococcus pneumoniae	ANP	A	500
3GPL	Exodeoxyribonuclease V, subunit RecD, putative	Deinococcus radiodurans R1	ANP	A	801
3GQI	Basic fibroblast growth factor receptor 1	Homo sapiens	ACP	A	775
3HQD	Kinesin-like protein KIF11	Homo sapiens	ANP	A	601
3I5X	ATP-dependent RNA helicase MSS116	Saccharomyces cerevisiae	ANP	A	1000
3IDB	cAMP-dependent protein kinase catalytic subunit alpha	Mus musculus	ANP	A	450
3IEV	GTP-binding protein era	Aquifex aeolicus	GNP	A	501
3JVV	Twitching mobility protein	Pseudomonas aeruginosa	ACP	A	400
3K5H	Phosphoribosyl-aminoimidazole carboxylase	Aspergillus clavatus	ATP	A	400
3Q86	Nucleoside diphosphate kinase	Staphylococcus aureus subsp. aureus	GTP	A	158
3QF7	Rad50	Thermotoga maritima	ANP	A	853
3QXC	Dethiobiotin synthetase	Helicobacter pylori	ATP	A	221
3RLF	Maltose/maltodextrin import ATP-binding protein MalK	Escherichia coli	ANP	A	2501
3RUV	Chaperonin	Methanococcus maripaludis	ANP	A	545
3RYC	Tubulin alpha chain	Ovis aries	GTP	A	600
3SUC	Preneck appendage protein	Bacillus phage phi29	ATP	A	856
3T1Q	Gliding protein mglA	Thermus thermophilus	GNP	A	197
3T9D	Inositol Pyrophosphate Kinase	Homo sapiens	ANP	A	401
3TII	Tubulin tyrosine ligase	Xenopus (Silurana) tropicalis	ANP	A	700
3UIE	Adenylyl-sulfate kinase 1, chloroplastic	Arabidopsis thaliana	ANP	A	402
3V2U	Galactose/lactose metabolism regulatory protein GAL80	Saccharomyces cerevisiae	ATP	C	523
4AZW	WBDD	Escherichia coli	ATP	A	1450
4DLU	GTPase HRas	Homo sapiens	GNP	A	201
4EHU	Activator of 2-hydroxyisocaproyl-CoA dehydratase	Clostridium difficile	ANP	A	302
4F6T	Molybdenum storage protein subunit alpha	Azotobacter vinelandii	ATP	A	301

Table S2. Ligands contained in Protein Data Bank structures used in study.

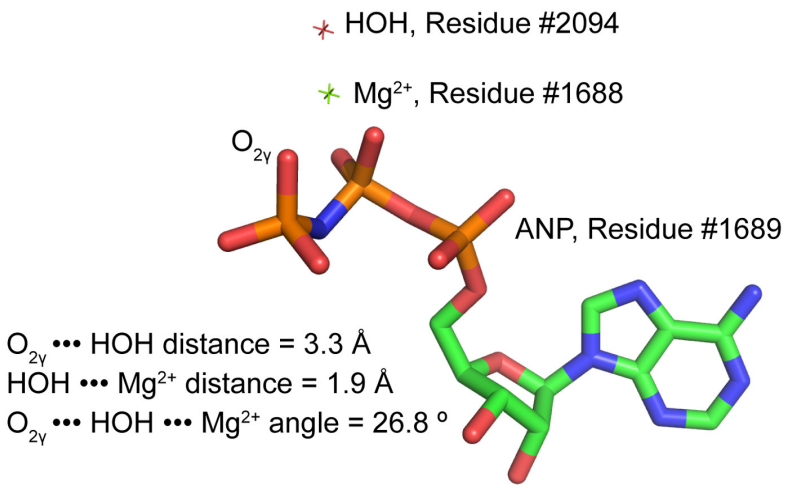
Nucleotide Analogs

Code	Ligand Name
ACP	PHOSPHOMETHYLPHOSPHONIC ACID ADENYLATE ESTER
ATF	PHOSPHODIFLUOROMETHYLPHOSPHONIC ACID-ADENYLATE ESTER
APC	DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER
ANP	PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER
SAP	ADENOSINE-5'-DIPHOSPHATE MONOTHIOPHOSPHATE
UTP	URIDINE 5'-TRIPHOSPHATE
CTP	CYTIDINE-5'-TRIPHOSPHATE
TTP	THYMIDINE-5'-TRIPHOSPHATE
GNP	PHOSPHOAMINOPHOSPHONIC ACID-GUANYLATE ESTER
DG3	2'-3'-DIDEOXYGUANOSINE-5'-TRIPHOSPHATE
GSP	5'-GUANOSINE-DIPHOSPHATE-MONOTHIOPHOSPHATE
GTP	GUANOSINE-5'-TRIPHOSPHATE
ATP	ADENOSINE-5'-TRIPHOSPHATE
DTP	2'-DEOXYADENOSINE 5'-TRIPHOSPHATE
AGS	PHOSPHOTHIOPHOSPHORIC ACID-ADENYLATE ESTER
AD9	ADP METAVANADATE
DCT	2',3'-DIDEOXYCYTIDINE 5'-TRIPHOSPHATE
G2P	PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER
DDS	2',3'-dideoxyadenosine triphosphate
DGT	2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE
DUT	DEOXYURIDINE-5'-TRIPHOSPHATE
UPC	2'-deoxy-5'-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]methyl}phosphoryl]uridine
QTP	3-{2-deoxy-5-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]oxy}phosphoryl]-beta-D-erythro-pentofuranosyl}-7-methyl-3H-imidazo[4,5-b]pyridine
D3T	2',3'-DIDEOXY-THYMIDINE-5'-TRIPHOSPHATE
23T	2',3'-DIDEOXY-THYMIDINE-5'-TRIPHOSPHATE
GTF	2'-deoxy-2',2'-difluorocytidine 5'-(tetrahydrogen triphosphate)
1GC	2'-deoxy-5'-O-[(R)-hydroxy{[(S)-hydroxy(phosphonooxy)phosphoryl]methyl}phosphoryl]guanosine
ZAN	5'-O-[(S)-hydroxy{[(S)-hydroxy(phosphonooxy)phosphoryl]amino}phosphoryl]adenosine
B44	N-(2-phenylethyl)adenosine 5'-(tetrahydrogen triphosphate)
GFH	2'-DEOXY-5'-O-[(R)-{[(R)-[(R)-FLUORO(PHOSPHONO)METHYL](HYDROXY)PHOSPHORYL]OXY}(HYDROXY)PHOSPHORYL]GUANOSINE
8DG	8-OXO-2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE
8GT	8-OXO-GUANOSINE-5'-TRIPHOSPHATE
3AT	3'-DEOXYADENOSINE-5'-TRIPHOSPHATE
CH1	3'-DEOXY-CYTIDINE-5'-TRIPHOSPHATE
DCP	2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE
MGT	7N-METHYL-8-HYDROGUANOSINE-5'-TRIPHOSPHATE
GCP	PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER
MGP	7-METHYL-GUANOSINE-5'-TRIPHOSPHATE

Table S3. Structure selection criteria.

The database was initially populated with all entries in the Protein Data Bank as of October 3rd, 2012 with resolution better than 2.7 Å that contained one of the ligands listed in Table S2. From 1866 structures, the following elimination criteria were applied, leaving 305 in a curated database.

1. No associated journal article was available
2. There was a mutation within 20Å of the active site ligand.
3. The authors indicated in the associated paper that a remote mutation affected the active site conformation.
4. Of the same protein, complexed with different ligands, only the highest resolution structure was retained.
5. The authors indicated that the structure was of a non-productive configuration.
6. Only structures containing the essential metal ion cofactors were retained for $O_{3\beta}$—P_{γ} and P_{α}—$O_{3\alpha}$ cleaving enzymes
7. If the ligand was in multiple conformations in the structure.
8. If any of the phosphoryl group atoms in the ligand had B-factors above 80 $Å^2$.
9. If any residues coordinating the ligand were absent.
10. If a protein catalyzed cleavage of multiple phosphoanhydride bonds.
11. Only enzymes that cleaved either the $O_{3\beta}$—P_{γ} or P_{α}—$O_{3\alpha}$ bonds or did not cleave any phosphoanhydride bond were used.
12. If an inhibitor was present that affected the active site conformation.



The role of additional interaction acceptance criteria is best understood through the specific example of the structure of the Focal Adhesion kinase (PDB ID = 2J0L). Chelation of the metal ion brings the solvent close enough to the phosphates to be considered as a direct interaction (O2γ ••• HOH distance = 3.3 Å). However, if the solvent is oriented such that the lone pairs interact with the metal, the implicitly placed hydrogens are not positioned to interact with the phosphate oxygens. Such interactions are excluded if both of the following additional criteria are true:

1) If the putative hydrogen bond donor (D) is closer to the metal ion (Me) than to the acceptor (O) on the ligand, specifically if the D•••O distance exceeds the Me•••D distance by more than 0.6 Å. (In this example D is the solvent oxygen, Me is the Mg2+ ion and O is the O2γ of the AMPPNP.

2) If the O•••D•••Me angle is less than 50°.

Figure S1. Criteria for excluding ion-coordinating atoms.

Table S4. Pairwise t-tests between structure sets.

Resolution Cutoff: 2.7Å, Water not included

		Non-bridging oxygens			Bridging oxygens		
Structure set	**# structures**	**Mean +/ Standard error**					
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites	155	2.4 ± 0.1	4.2 ± 0.2	4.6 ± 0.2	0.1 ± 0.0	0.3 ± 0.0	0.8 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	100	2.7 ± 0.1	3.3 ± 0.1	4.2 ± 0.2	0.1 ± 0.0	0.4 ± 0.1	0.4 ± 0.1
Non-catalytic binding sites	50	2.0 ± 0.2	2.5 ± 0.2	4.1 ± 0.3	0.2 ± 0.1	0.3 ± 0.1	0.5 ± 0.1
p-values from pairwise t-tests							
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		0.09	6.3E-10	0.07	0.30	0.78	0.01
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.01	0.01	0.72	0.55	0.15	0.27
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.24	3.4E-05	0.07	0.55	0.02	5.0E-07

Resolution Cutoff: 2.0Å, Water included

		Non-bridging oxygens			Bridging oxygens		
Structure set	**# structures**	**Mean +/ Standard error**					
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites	72	4.2 ± 0.2	5.3 ± 0.2	7.0 ± 0.3	0.1 ± 0.1	0.3 ± 0.1	1.1 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	45	4.7 ± 0.2	4.9 ± 0.2	7.1 ± 0.3	0.4 ± 0.1	0.7 ± 0.1	0.6 ± 0.1
Non-catalytic binding sites	17	3.8 ± 0.3	4.2 ± 0.4	7.2 ± 0.4	0.0 ± 0.0	0.4 ± 0.1	0.4 ± 0.1
p-values from pairwise t-tests							
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		0.58	0.04	0.78	0.50	0.78	2.9E-04
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.12	0.24	0.91	0.06	0.16	0.38
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.15	0.28	0.80	0.07	0.01	2.0E-04

Resolution Cutoff: 2.0Å, Water not included

		Non-bridging oxygens			Bridging oxygens		
Structure set	**# structures**	**Mean +/ Standard error**					
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites	72	2.5 ± 0.1	4.4 ± 0.2	4.8 ± 0.2	0.1 ± 0.1	0.2 ± 0.1	0.8 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	45	2.8 ± 0.2	3.6 ± 0.2	4.4 ± 0.3	0.1 ± 0.1	0.5 ± 0.1	0.4 ± 0.1
Non-catalytic binding sites	17	2.0 ± 0.2	3.0 ± 0.4	4.9 ± 0.4	0.0 ± 0.0	0.2 ± 0.1	0.4 ± 0.1
p-values from pairwise t-tests							
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		0.30	2.6E-03	0.76	0.52	0.74	0.01
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.10	0.25	0.30	0.53	0.19	1.00
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.30	0.02	0.30	1.00	0.01	2.2E-04

Blue shading indicates a p-value ≤ 0.05

Table S5. Pairwise t-tests between stucture sets that contain riding hydrogens. The interpretation in the main paper is based on Table S4, without riding hydrogens.

Resolution Cutoff: 2.7Å, Water not included

		Non-bridging oxygens			Bridging oxygens		
Structure set	**# structures**	**Mean +/ Standard error**					
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites	155	2.2 ± 0.1	3.8 ± 0.1	4.5 ± 0.2	0.1 ± 0.0	0.2 ± 0.0	0.7 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	100	2.6 ± 0.1	3.2 ± 0.2	4.0 ± 0.2	0.1 ± 0.0	0.4 ± 0.1	0.4 ± 0.1
Non-catalytic binding sites	50	1.9 ± 0.2	2.3 ± 0.2	3.9 ± 0.3	0.2 ± 0.1	0.2 ± 0.1	0.5 ± 0.1
p-values from pairwise t-tests							
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		0.23	1.1E-08	0.03	0.28	0.90	0.04
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.01	1.7E-03	0.70	0.80	0.17	0.23
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.08	3.7E-03	0.03	0.31	0.04	1.4E-05

Resolution Cutoff: 2.0Å, Water included

		Non-bridging oxygens			Bridging oxygens		
Structure set	**# structures**	**Mean +/ Standard error**					
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites	72	3.9 ± 0.1	4.8 ± 0.2	6.7 ± 0.3	0.1 ± 0.1	0.3 ± 0.1	1.0 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	45	4.6 ± 0.2	4.8 ± 0.2	6.8 ± 0.3	0.4 ± 0.1	0.6 ± 0.1	0.6 ± 0.1
Non-catalytic binding sites	17	3.8 ± 0.4	4.2 ± 0.4	7.0 ± 0.4	0.0 ± 0.0	0.3 ± 0.1	0.4 ± 0.1
p-values from pairwise t-tests							
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		0.86	0.20	0.64	0.53	0.91	5.1E-04
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.13	0.21	0.80	0.07	0.19	0.26
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.04	0.97	0.82	0.07	0.02	1.1E-03

Resolution Cutoff: 2.0Å, Water not included

		Non-bridging oxygens			Bridging oxygens		
Structure set	**# structures**	**Mean +/ Standard error**					
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites	72	2.2 ± 0.1	4.0 ± 0.2	4.5 ± 0.2	0.1 ± 0.1	0.2 ± 0.0	0.7 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	45	2.7 ± 0.2	3.6 ± 0.2	4.2 ± 0.3	0.1 ± 0.1	0.4 ± 0.1	0.3 ± 0.1
Non-catalytic binding sites	17	2.0 ± 0.2	3.0 ± 0.4	4.7 ± 0.3	0.0 ± 0.0	0.2 ± 0.1	0.3 ± 0.1
p-values from pairwise t-tests							
		α	**β**	**γ**	**O5'**	**α**	**β**
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		0.55	0.02	0.62	0.60	0.90	0.01
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.09	0.17	0.18	0.60	0.22	0.82
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.09	0.21	0.18	0.92	0.03	1.6E-03

Blue shading indicates a p-value ≤ 0.05

Table S6. Structures in which the $O_{3\beta}$ oxygen of a nucleoside triphosphate has interactions other than water. Conserved motifs in protein sequence data were identified using ProSite.

* From paper referencing PDB structure.
** From PROSITE

		Residue				Motif			
PDB ID	**Enzyme**	**Atom Name**	**Type**	**Number**	**Chain**	**AA Range**	**Motif sequence**	**Position in Motif**	**Name of Motif/Region**
1A5U	Pyruvate kinase	ND2	ASN	74	A				
		NH2	ARG	119	A				
1A82	Dethiobiotin synthetase	N	GLU	12	A	8 - 16:	GtdtevGKT	5	Similarity to Walker A motif
		OE1	GLU	12	A				
1AM4	Cdc42Hs G-protein, P50-RHOGAP	N	ALA	513	D	510 - 517:	GdgavGKT	4	Walker A motif
1AQ2	Phosphoenolpyruvate carboxykinase	N	GLY	251	A	248-255	GlsgtGKT	4	Walker A motif
1B63	MUTL (DNA mismatch repair ATPase)	N	PHE	94	A	93 - 99:	GFRGEAL	2	DNA mismatch repair proteins mutL / hexB / PMS1 signature**
		N	GLY	96	A			4	
1BIF	6-Phosphofructo-2-Kinase/Fructose-2,6-Bisphosphatase Bifunctional enzyme	N	ALA	48	A	45-52	GlparGKT	4	Walker A motif
		NZ	LYS	172	A				
1BXR	Carbamoyl Phosphate Synthetase	NE2	HIS	243	A				
1C1Y	Rap1A, c-Raf1 serine/threonine kinase	N	GLY	13	A	10 - 17:	GsggvGKS	4	Walker A motif
1E4G	FTSA Cell division protein	N	SER	17	T				
		OG	SER	17	T				
1F2U	Rad50 ABC-ATPase	N	GLY	33	A	30 - 37:	GqngsGKS	4	Walker A motif
1G17	SEC4 G-protein	N	GLY	30	A	27 - 34:	GdsgvGKS	4	Walker A motif
1G7T	IF2, EIF5B initiation factor	OD1	ASP	15	A				
1GJV	a-ketoacid dehydrogenase kinase	N	GLY	337	A	238 - 374			Histidine kinase domain profile**
1GS5	N-Acetyl-L-Glutamate kinase	N	GLY	11	A				
1I6I	KIF1A Motor domain	N	GLY	100	A	97 - 104:	GqtgaGKS	4	Walker A motif
1JST	Cyclin-dependent kinase-2	O	GLY	13	A	10-33			Protein kinases ATP-binding region signature**
1K5D	Ran G-protein	N	GLY	20	A	17 - 24:	GdggtGKT	4	Walker A motif
1KO5	Gluconate kinase	N	GLY	18	A	15-22	GvsgsGKS	4	Walker A motif
1LII	Adenosine kinase	N	ALA	316	A	312 - 325:	DTnGAGDafvGGFL	5	pfkB family of carbohydrate kinases signature 2**
		N	GLY	317	A			6	
		O	THR	313	A			7	
1M2O	Sar1, Sec23 G-protein and GAP	N	ASN	33	B	30 - 37:	GldnaGKT	4	Walker A motif
		NH1	ARG	722	A				
1MH1	Rac1 G-protein	N	ALA	13	A	10 - 17:	GdgavGKT	4	Walker A motif
1MX0	topoisomerase subunit	N	GLY	109	A				
1N6H	Rab5a G-protein	N	ALA	30	A	27 - 34:	GesavGKS	4	Walker A motif

1NGJ	Heat-Shock 70 protein	N	GLY	202	A	197 - 210:	IFDLGGGTf dvSIL	6	Heat shock hsp70 proteins family signature 2**
1NHI	MUTL (DNA mismatch repair ATPase)	N	PHE	94	A	93 - 99:	GFRGEAL	2	DNA mismatch repair proteins mutL / hexB / PMS1 signature**
		N	GLY	96	A			4	
1OXV	GlcV ABC-ATPase	N	GLY	41	A	38 - 45:	GpsgaGKT	4	Walker A motif
1PVG	ATPase region of topoisomerase II	N	ARG	141	A				
		N	GLY	143	A				
1R2Q	Rab5a GTPase Domain	N	ALA	30	A	27 - 34:	GesavGKS	4	Walker A motif
1SVM	SV40 large T antigen helicase domain	N	ASP	429	A	426 - 433:	GpidsGKT	4	Walker A motif
1SX3	GroEL14 chaperonin protein	N	THR	89	A				
		N	THR	90	A				
1TZ6	Aminoimidazole Riboside kinase	NZ	LYS	187	A				
1W2C	Inositol (1,4,5) Triphosphate 3-kinase	OD2	ASP	416	A				$ID_{416}FG$ motif*
1W5A	FTSZ Dimer tubulin homolog	N	GLY	134	A	123 - 144:	DMvFITCGl GGGTGTGS APVVA	12	FtsZ protein signature 2**
		N	THR	135	A			13	
1W5T	ORC2 DNA Replication Initiation protein	N	GLY	62	A	59 - 66:	GrvgiGKT	4	Walker A motif
		NH1	ARG	260	A				
1XDP	Polyphosphate kinase	NE2	HIS	592	A				
1Y8P	Kinase 3, lipoyl domain 2 of pyruvate dehydrogenase complex	N	GLY	323	A	238 - 362			Histidine kinase domain profile**
1YTM	phosphoenolpyruvate carboxykinase	N	GLY	245	A	242 - 249:	GlsgtGKT	4	Walker A motif
1Z2P	Inositol 1,3,4-trisphosphate 5/6-kinase	ND2	ASN	210	X				
1ZXM	Topo IIa ATPase	N	ARG	162	A				
		N	GLY	164	A				
2AJ4	Galactokinase	OG	SER	171	A				
2BMU	UMP kinase	OG	SER	10	A				
2DB3	DEAD-box RNA helicase	N	GLY	292	A	276 - 453:		16	Superfamilies 1 and 2 helicase ATP-binding type-1 domain profile**
		NH2	ARG	582	A	477 - 621		105	Superfamilies 1 and 2 helicase C-terminal domain profile**
2DWP	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase	N	ALA	44	A	41- 48:	GlparGKT	4	Walker A motif
2EWS	Pantothenate kinase	OG1	THR	99	A				pseudo p-loop*
2FPM	RadA recombinase	N	GLY	108	A	105 - 112:	GvfgsGKT	4	Walker A motif
2GJK	Upf1 helicase core	N	GLY	495	A	492 - 499:	GppgtGKT	4	Walker A motif
		NH2	ARG	703	A				

2IO7	Bifunctional glutathionylspermidine synthetase/amidase	OE2	GLU	330	A				Highly conserved residue*
2IS4	UvrD DNA helicase	N	GLY	32	A	8 - 286		24	UvrD-like DNA helicase ATP-binding domain profile**
		NH2	ARG	284	A			276	
2J4J	Uridylate kinase	NZ	LYS	10	D				
2JLV	NS3 Helicase	N	GLY	196	A	180 - 336		16	Superfamilies 1 and 2 helicase ATP-binding type-1 domain profile**
		NH1	ARG	463	A	331 - 512		132	Superfamilies 1 and 2 helicase C-terminal domain profile**
2P0C	MER Tyrosine-protein kinase	OD1	ASN	728	A	592 - 858			Protein kinase domain profile**
2PZE	CFTR first nucleotide binding domain	N	GLY	461	A	458 - 465:	GstgaGKT	4	Walker A motif
2QO9	EphA3 kinase	O	ALA	629	A	627 - 653:			Protein kinases ATP-binding region signature**
2QQ0	Thymidine kinase	N	TYR	13	A	10 - 17:	GpmysGKT	4	Walker A motif
2QT0	Nicotinamide riboside kinase 1	NH1	ARG	132	A				
2QTV	Sec23, Sar1, Sec31 G-protein and GAP proteins	N	ASN	33	B	30 - 37:	GldnaGKT	4	Walker A motif
		NH1	ARG	722	A				
2W00	HSDR Subunit OF ECOR124I restriction enzyme	NH2	ARG	691	A				
2X9H	Myosin-2 motor domain	N	GLY	182	A	179 - 186:	GesgaGKT	4	Walker A motif
		ND2	ASN	233	A				
3A1C	CopA ATPase	OG1	THR	426	A				
3A4L	O-phosphoseryl-tRNA(Sec) kinase	N	GLY	14	A	11 - 18:	GlpgvGKS	4	Walker A motif
		NH1	ARG	120	A				Rx3R motif*
3DCB	NOD kinesin	N	GLY	90	A	87 - 94:	GqtgtGKS	4	Walker A motif
3DT7	Phosphoenolpyruvate carboxykinase	N	ALA	287	A	284 - 292:	FPSACGKTN	4	Phosphoenolpyruvate carboxykinase (GTP) signature**
3EHG	DesK histidine kinase	ND1	HIS	335	A				
3EW9	RADA recombinase	N	GLY	108	A	105 - 112:	GmfgsGKT	4	Walker A motif
3FHT	DEAD-box protein 5 RNA helicase	N	GLY	141	A	125 - 295		16	Superfamilies 1 and 2 helicase ATP-binding type-1 domain profile**
		NH2	ARG	432	A	306 - 467		126	Superfamilies 1 and 2 helicase C-terminal domain profile**
3GON	Phosphomevalonate kinase	OG	SER	213	A				
3GPL	RecD2 helicase	N	GLY	363	A	359 - 367:	GgpgtGKS	4	Walker A motif
		NH1	ARG	493	A				
3HQD	Kinesin Eg5 motor domain	N	GLY	108	A	105 - 112:	GqtgtGKT	4	Walker A motif
		ND2	ASN	229	A	17 - 290		212	Kinesin motor domain profile

3I5X	DEAD box protein Mss116p helicase	N	GLY	155	A	139 - 326		16	Superfamilies 1 and 2 helicase ATP-binding type-1 domain profile**
		NH2	ARG	469	A	355 - 512		114	Superfamilies 1 and 2 helicase C-terminal domain profile**
3IEV	ERA G-protein	N	ASN	13	A	10 - 17:	GkpnvGKS	4	Walker A motif
3QF7	Mre11, Rad50 nuclease and ABC ATPase	N	GLY	33	A	30 - 37:	GpngaGKS	4	Walker A motif
3RLF	Maltose-binding protein,maltose transporter	N	GLY	39	A	36 - 43:	GpsgcGKS	4	Walker A motif
3RUV	Cpn-rls chaperonin	N	THR	94	A				
3RYC	Tubulin, RB3 stathmin-like domain	N	GLY	144	A	142 - 148:	GGGTGSG	3	Tubulin subunits alpha, beta, and gamma signature**
		N	THR	145	A			4	
3T1Q	MglA, MglB G-protein and GAP	N	LEU	22	A	19 - 26:	GpglsGKT	4	Walker A motif
3UIE	Adenosine 5'-phosphosulfate kinase	N	GLY	111	A	108 - 115:	GlsgsGKS	4	Walker A motif
4DLU	H-Ras G-protein	N	GLY	13	A	10 - 17:	GaggvGKS	4	Walker A motif
4EHU	Activator of 2-hydroxyisocaproyl-CoA dehydratase (ATPase)	N	SER	10	A				
		N	GLY	102	A				
		OG	SER	10	A				

Table S7. Hydrogen bonds between positively charged donors and γ-oxygens. The resolution cutoff is 2.7 Å.

Structure set	**# structures**	**Mean +/ Standard error**
$O_{3\beta}$—P_{γ}-cleaving sites	155	1.3 ± 0.1
P_{α}—$O_{3\alpha}$-cleaving sites	100	1.6 ± 0.1
Non-catalytic binding sites	50	2.0 ± 0.2
p-values from pairwise t-tests		
$O_{3\beta}$—P_{γ}-cleaving sites vs. non-catalytic binding sites		4.6E-03
P_{α}—$O_{3\alpha}$-cleaving sites vs. non-catalytic binding sites		0.13
$O_{3\beta}$—P_{γ}-cleaving sites vs. P_{α}—$O_{3\alpha}$-cleaving sites		0.09

Table S8. Pairwise t-tests between oxygen types. The resolution cutoff is 2.0 Å and water is included in the analysis.

	Non-bridging oxygens			Bridging oxygens		
	α x β	**α x γ**	**β x γ**	**O5' x α**	**O5' x β**	**α x β**
$O_{3\beta}$—P_{γ}-cleaving sites	1.4E-04	0.56	2.4E-03	0.22	< 2E-16	2.2E-12
P_{α}—$O_{3\alpha}$-cleaving sites	0.70	0.56	0.33	0.11	0.21	0.76
Non-catalytic binding sites	0.58	0.37	0.70	0.18	0.11	0.82

CONCLUSIONS

The broad goal of the research described here, was to characterize the factors that contribute to catalysis of the forming and breaking of high energy bonds such as those in ATP. Previous work in the laboratory had established that the anomeric effect explains much of the chemistry of high energy bonds, filling in shortfalls in text-book explanations. The specific question addressed here was whether enzymes harness the anomeric effect for catalytic gain. For this to be true, the anomeric effect had to be amenable to modulation through interaction with other molecules. The potential was there; profuse electron density surrounding phosphoryl oxygens is the source of the exceptional anomeric effect in molecules like ATP. If enzymes could exploit this unique chemistry, through the anomeric effect they could drive scissile bond weakening, a critical step in the catalytic reaction. Initial indications were that perturbing the electron density surrounding the phosphoryl oxygens would lead to changes in the anomeric effect and scissile bond length: tying up electron density through protonation decreased the anomeric effect.[15] Initial discoveries from our lab and others inspired several lines of research that were pursued in this dissertation; Firstly how might enzymes influence the anomeric effect for catalytic gain; what configurations of active sites enhance the anomeric effect? Secondly is the anomeric effect a common factor in the catalysis of high energy bonds? Finally, can the anomeric effect account for substantial rate-enhancement? All of these questions were aimed at establishing whether the anomeric effect is a major, broadly applicable, component of enzyme-catalyzed phosphoryl transfer. Below, is a

summary of progress made in this dissertation. The final section "Future Directions" details, of the remaining questions, how the most pressing would best be addressed.

What configurations of enzyme active sites enhance the anomeric effect? Using high level computation, this dissertation demonstrates that active sites can have a significant impact on scissile bond length/strength. Hydrogen bonds at the γ-phosphoryl group reduce scissile bond length and hydrogen bonds at the β-phosphoryl, particularly at the bridging β-oxygen, increase scissile bond length. The water interactions investigated in the early parts of this research and the interactions from an amino acid backbone analog in the latter part serve as a model for general polar interactions in an enzyme active site. This research is the first to detail how placement of hydrogen bonds relative to the phosphoryl groups in a nucleotide substrate can alter the anomeric effect.

The work here nicely details the effects of enzyme-substrate interactions on substrate chemistry for ATP-type substrates; however, how environment impacts other high energy substrates is unknown. If similarities with ATP-type substrates are discovered, it would suggest the anomeric effect is broadly associated with catalysis of high energy bonds, prompting reevaluation of many catalytic mechanisms.

Is the anomeric effect a common catalytic mechanism? In the second part of this dissertation, computational studies were complemented with a database search where analysis of substrate-bound structures of phosphoryl transfer enzymes from the Protein Data Bank showed that, on average, enzyme active sites are configured to favor substrate-enzyme interactions that enhance anomeric bond destabilization. Results support the importance of the anomeric effect in a broad range of enzyme that transfer phosphate groups.

Here the focus was on high energy bonds in nucleoside triphosphates. Targeting enzymes that bind triphosphates was a logical first step for documenting the anomeric effect in nature; nucleotide-binding enzymes are ubiquitious, providing large datasets. However, it is unknown whether the anomeric effect is important for catalysis of Nature's other high energy substrates. An analysis of the impact of interactions on the anomeric effect in other substrates might necessitate a different bioinformatics strategy then was used here. A sufficient dataset size will require a more generalized approach with the study group composed of enzymes that cleave any high energy bond and the control, enzymes that cleave non-high energy bonds. It won't be a trivial project but important, if realized, as it will demonstrate a fundamental link between catalysis of high energy bonds and the anomeric effect.

To what degree does the anomeric effect enhance rate? This work describes a first step toward demonstrating, experimentally, a relationship between the anomeric effect and rate. The kinetic analysis of our AK mutant exhibits a 100-fold decrease in kcat relative to wild type. Small molecule quantum studies suggest such a mutation would have a large impact on the anomeric effect and bond length. However an accurate assessment of the impact on rate will require calculation in the context of the full AK enzyme. The work here has established proof of concept that perturbing residues predicted by computation to have an impact on rate does indeed result in an experimentally observed rate change. However additional experiments are needed to precisely correlate the anomeric effect with rate. Another major unanswered question is how much of the total rate-enhancement from an enzyme can be attributed to the anomeric effect relative to other catalytic factors.

Priority in addressing remaining questions. The major remaining questions fall broadly into two categories; 1) does the anomeric effect influence catalysis of other high energy bonds? and 2) for the high energy O—P bonds studied here, what is the relationship between the anomeric effect and rate? Of these two broad categories of interest, determining the impact of rate is both the most pressing question and the most difficult to address. Establishing the ability of the anomeric effect to impact reaction rate is crucial for the acceptance by the scientific community that the anomeric effect is as a key catalytic factor in phosphoryl transfer. The future directions presented below will outline the key components of this question and lay out a strategy for answering it.

FUTURE DIRECTIONS

Overall Rationale:

Our group has previously demonstrated that the anomeric effect is important in the ground state energetics of biologically important "high energy" phosphate compounds. We have shown that the environment affects the strength of these interactions, and that active sites are, on average, configured to modulate the anomeric effect to destabilize bonds broken during the reaction cycle. This research suggested here tests the hypothesis that enzymes do, in fact, modulate such anomeric interactions for catalytic effect.

The hypothesis tested will involve detailed studies of one model enzyme, to dissect the contributions of active site elements to the anomeric effect through calculations combined with experimental structure and kinetics of active site mutants. The labor required will restrict studies to a single example enzyme, but the results should

be generalizable to some extent in providing a first quantitative indication of the importance of anomeric bond destabilization in the overall catalytic picture.

This research requires several steps be performed in succession. First, atomic coordinates for the enzyme-substrate complex in key states along the reaction path must be determined either experimentally or computationally. These coordinates are then used to map the AK reaction path from ground to transition state in preparation for calculation of reaction rate. All of these calculations will first be performed in the wild type enzyme and then repeated in mutants of arginine kinase. It remains difficult to perform quantum mechanical calculations with accuracy sufficient to predict rate. The plan is to examine differences between wild-type and perturbed forms such that known shortcomings / systematic errors in the calculations will cancel out, making it easier to establish correlation between the anomeric effect and rate.

Aim: Determine the extent to which the anomeric effect contributes to rate-enhancement in arginine kinase.

Our lab has demonstrated that an anomeric component is an important part of high energy bonds that can be perturbed by the environment and that nucleotide-utilizing enzymes are, on average, configured to enhance the anomeric effect above the solution-state level. These analyses were performed on models of isolated substrates. A logical next step is to ask to what extent an enzyme can manipulate the local environment and thereby modulate the anomeric effect. There are no known experimental methods for quantitatively dissecting anomeric effects from other components to an enzyme mechanism. The approach proposed uses computer calculations to do this – computer calculations that will be cross-checked comparing the calculated and experimentally

measured effects of perturbations to the active site. The quantitative approach to be used is computation and labor-intensive, so is performed with a single model system that is well characterized structurally and enzymologically.

Step 1: Define coordinates for key stuctures in the arginine kinase reaction path.

For both calculation of rate and for an assessment of impact of the anomeric effect, computational mapping of the reaction path is required. Atomic coordinates from protein crystal structures are often used as input data for molecular modeling studies. For our planned calculation of the AK reaction path, it is useful to have initial coordinates for key steps in the reaction, including the ground state Michaelis structure, the product structure and a transition state structure. Thus the initial phase of this work concerns acquisition of relevant structures.

Initial coordinates for modeling of the transition state will be taken from the previously published crystal structure mimicking the transition state of the AK reaction determined at a particularly high resolution (1.2 Å)[52] with a coordinate precision of ± 0.04 Å. There is a structure for the substrate-free form but there are currently no available crystal structures of the ground state Michaelis complex or the product complex. However, as detailed below, I have made progress toward determination of these structures in my dissertation work.

Progress toward determining the Michaelis and product complex crystal structures. – Multiple crystallization experiments have been performed in an attempt to attain the Michaelis, product, and several binary complexes of arginine kinase. Hundreds of crystallization conditions were tested for each target complex with the use of commercially available screens including the Wiz I&II, Wiz III&IV screen from Emerald

Biosystems and the Hampton HT screen. Each screen contains 96 different crystallization conditions. Multiple parameters were tested in these trials including the incubation temperature of the crystal and the ratio of the volume of the protein drop solution to the volume of well solution. The different crystal incubation temperatures tested were 4°C and 20°C and the three different ratios of drop to well solution were tested; 1:1, 2:1 and 1:2 drop:well respectively.

To capture the Michaelis complex where all ligands are bound, our strategy was to trap the enzyme in the pre-reaction state using an unreactive analog of either the ATP or arginine substrate. Four different analogs of ATP were used as well as two analogs of arginine (Table 3). Several datasets were collected that resulted in structures that were roughly equivalent to the already-published transition state analog structure. In the published structure there is a planar nitrate molecule in the active site centered roughly between the ADP and arginine substrates.[52] In datasets #11 and #9 (Table 4), there was also a nitrate present at that location, likely due to nitrate in the reaction buffer. It was found that the commercially available ATP analogs of purity < 95% resulted in crystals containing ADP in the active site. As binding of ADP in the active site was not observed with ATP analogs of > 95% purity, it is likely the ADP binding is due to impurities in the analog sample and not the result of hydrolysis of the analog or of phosphosphoryl transfer by arginine kinase.

Several attempts at collecting a binary complex and a product complex were also made. Binary complexes with an ATP bound and in the closed, reaction-ready conformation, could also aid us in understanding of how the ATP binds in the active site. To produce the product complex our intent was to create an analog of the

phosphoarginine product by mixing either sodium metavanadate ($NaVO_3$) or Sodium orthovanadate (NO_3VO_4) with arginine. Vanadate ions have been used to form analogs of phosphorylated substrates.[132,133]

Table 3. Ligands used in attempted crystallization of complexes of arginine kinase.

AK variant	Target complex	Ligands			
WT		PNP			
WT		PCP			
WT		ATPγS			
WT		ATP			
WT	Binary	PNP	Mg^{2+}		
WT	Binary	PCP	Mg^{2+}		
WT	Binary	ATPγS	Mg^{2+}		
WT	Binary	ATP	Mg^{2+}		
WT	Binary	Arg	Mg^{2+}		
WT	Michaelis	PNP	Mg^{2+}	Arg	
WT	Michaelis	PCP	Mg^{2+}	Arg	
WT	Michaelis	ATPγS	Mg^{2+}	Arg	
WT	Michaelis	ATP	Mg^{2+}	Iminoethyl-L-ornithine	
WT	Michaelis	ATP	Mg^{2+}	L-ornithine	
WT	Michaelis	ADP	Mg^{2+}	Arg	BeF_3
WT	Product	ADP	Mg^{2+}	Arg	$NaVO_3$
WT	Product	ADP	Mg^{2+}	Arg	Na_3VO_4
R280K	Michaelis	PNP	Mg^{2+}	Arg	
R280K	Michaelis	PCP	Mg^{2+}	Arg	
R280K	TSA	ADP	Mg^{2+}	Arg	NO_3

Footnote 1. WT is wild type; R280K is the mutant AK; TSA is transition state analog; PNP is Adenosine 5'-(β,γ-imido)triphosphate; PCP is β,γ-Methyleneadenosine 5'-triphosphate and ATPγS is Adenosine 5'-[γ-thio]triphosphate.

Preliminary structure solutions to several of the datasets were achieved (Table 4). Datasets 1, 2, 3 and 5 were shown to be very similar in terms of global protein

conformation to the already published substrate-free form of arginine kinase[134] and datasets 9 and 11 were shown to be similar in terms of global protein conformation and location of active site substrates to the published transition state analog structure.[52] Furthermore, a preliminary solution for datasets 15 and 16 indicated that the ATP analog (PNP) was bound in a non-productive conformation in the active site. At this stage after several failed attempts at attaining the desired complexes, priorities were shifted to other aspects of the project that were deemed more likely to succeed. However several datasets remain unsolved and it is possible that one of these may yield the desired structure.

New commercial screens from Microlytic became available after our crystallization experiments ended. Future experiments could make use of a these new screens. Alternatively there are several other commercially available additive or salt screens that may be of use such as the Hampton Additive Screen HT or the Hampton SaltRX.

Initial coordinates for structures along the arginine kinase reaction path can be generated using computation. Experience in the lab shows that the availability of a high resolution crystal structure with high coordinate precision is an asset in attaining an energy minimum that remains consistent with experimental data. However, failing acquisition of a Michaelis or product complex, an alternative solution is to infer starting coordinates by computationally substituting ATP into the active site of the model attained from transition state analog crystal structure and then performing an energy minimization.

Table 4. Diffraction datasets of arginine kinase-substrate complexes. An entry of "Apo" or "TSA" in the Result column indicates that a preliminary structure solution shows the dataset structure is similar in conformation and location of any ligands to the previously published substrate-free[134] and transition state analog[52] arginine kinase structures.

Dataset number	Target complex	Ligands	Resolution	Result
1	Binary	PCP	3	Apo
2	Binary	PCP	1.9	Apo
3	Binary	PNP	2	Apo
4	Binary	PNP	2.7	
5	Binary	PNP	2.7	Apo
6	Binary	PNP	3	
7	Michaelis	ATPγS Mg^{2+} Arg	4	
8	Michaelis	ATPγS Mg^{2+} Arg	2	
9	Michaelis	ATPγS Mg^{2+} Arg	1.7	TSA
10	Michaelis	ATPγS Mg^{2+} Arg	1.5	
11	Michaelis	ATPγS Mg^{2+} Arg	1.4	TSA
12	Michaelis	ATPγS Mg^{2+} Arg	2.3	
13	Michaelis	ATPγS Mg^{2+} Arg	2.07	
14	Michaelis	ATPγS Mg^{2+} Arg	1.8	
15	Michaelis	PNP Mg^{2+} Arg	2.1	
16	Michaelis	PNP Mg^{2+} Arg	1.9	
17	Product	ATP Mg^{2+} Arg $NaVO_3$	2.8	

Footnote 2. WT is wild type; R280K is the mutant AK; TSA is transition state analog; Apo is the substrate-free arginine kinase structure; PNP is Adenosine 5'-(β,γ-imido)triphosphate; PCP is β,γ-Methyleneadenosine 5'-triphosphate and ATPγS is Adenosine 5'-[γ-thio]triphosphate.

Coordinates determined in Experiment 1 are used as input for Experiment 2, calculation of the arginine kinase reaction path.

Step 2: Modeling of the arginine kinase reaction path.

Calculation of rate necessitates a complete model of the reaction path[135] as the full path helps to differentiate the true transition state from other saddle points along the potential energy surface. Thus our next step is to calculate the full arginine kinase reaction path using a full quantum mechanics/molecular mechanics (QM/MM) environment model.

Calculations of the reaction path will be performed using a model of the full enzyme where high level quantum mechanics is interlaced with lower level molecular mechanics applied to more distant regions in a hybrid approach termed "QM/MM". Here, ligands and active site amino acids and waters are treated with the more accurate model chemistry, while the outer protein shell modeled with a simpler force field, endows the simulation with a more realistic electrostatic environment. Although relatively new, QM/MM is becoming state-of-the art for computational treatment of enzymes.

In general phosphoryl transfer reactions can adopt a fully dissociative reaction path, fully associative or an intermediate mechanism (Figure 30). In an associative transition state, the transferring phosphate is covalently bound to both the phosphate donor and acceptor, forming a pentacovalent intermediate. For a dissociative mechanism, a PO_3 metaphosphate intermediate is fully dissociated from both the donor and acceptor molecules. Creatine kinase, a close homolog of arginine kinase, has been shown to follow a partially associative reaction path.[136] It is generally believed that the AK reaction reaction path is similar[137], although the calculation has never been done. Thus, modeling

of the reaction rates will have to start with rigorous examination of each of the plausible transition states.

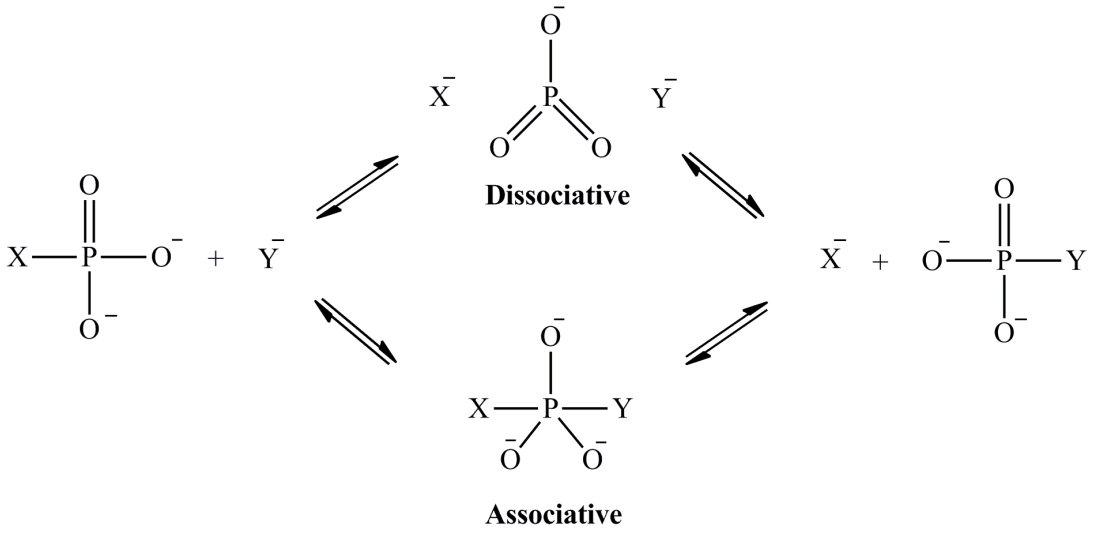


Figure 30. Transition states can be dissociative, associative or a hybrid of the two. In the dissociative transition state, there are no bonds to the leaving group or phosphor-acceptor, giving rise to a metaphosphate intermediate.

The reaction path requires definition of a reaction coordinate, generally a scalar function of key atomic coordinates (perhaps involved in formation or breakage of a covalent bond) that vary unimodally with the reaction. The reaction path is then the full set of reactant, solvent and enzyme atomic coordinates, energy optimized at each point along the reaction coordinate.

Calculating rate. Once the correct ground and transition state structures are known, their respective enthalpies can be calculated and thus the change in enthalpy. Extensive free energy calculations with computationally costly molecular dynamics sampling of structures orthogonal to the reaction coordinate will be required to generate an estimate of entropy. With estimates of the change in enathalpy and entropy in hand $\Delta G^{\circ\ddagger}$ can be calculated. By transition state theory the rate is:

$$k(T) = \frac{k_B T}{h} e^{-\Delta^{\ddagger}G^{\circ}/RT},$$

where k is the rate at Temperature, T. k_B is the Boltzmann constant, h is Plank's constant, $\Delta^{\ddagger}G^{\circ}$ is Gibb's free energy of activation and R is the gas constant.

Step 3: Comparison of experimental kinetics to calculated anomeric effects in site-directed mutants.

The final ground and transition state models produced in Step 2 will be used for an assessment of which residues in the AK active site most influence the anomeric effect. These residues will be targeted for mutation both experimentally and *in silico*.

Designing mutants – Our hypothesis is that the anomeric effect is a large contributor to rate enhancement relative to other factors that may be involved in the AK mechanism. Thus, while mutations must target the anomeric effect, the experimental design must also take into account the potential impact of other factors on rate. Each facet of catalytic mechanism investigated will require a small panel of mutants tailored primarily to perturb that factor. Thus time and resource limitations restrain our search to the most likely factors contributing to catalysis in AK. The current (limited) understanding of the AK catalytic mechanism places electrostatic stabilization of the transition state as a probable factor.[8] Electrostatic stabilization has in fact been trumpeted as the major means of catalytic rate enhancement of phosphoryl transfer and hydrolysis by several influential groups.[46,138] It follows that any investigation of the factors responsible for enzyme rate enhancement must be concerned with the potential contribution of electrostatics.

Once the reaction path is known and mutations are designed that target either primarily the anomeric effect or primarily electrostatics, mutations will be made *in silico*

and the reaction path recalculated to determine the perturbations from each mutant to rate. Strong correlations between changes in experimental rates and changes in computationally determined rates will serve as an indicator that our calculations are correct.

Step 4: Comparison of calculated and measured perturbations to the anomeric effect and electrostatics in AK mutants.

AK mutants will be engineered, expressed, purified and characterized kinetically by methods that are already well established in this laboratory.[115] Computed magnitudes of the anomeric effect and of the electrostatics will be compared to experimental rates of reaction for evidence that either factor is correlated with reaction rates.

Part a: Partial activity mutants are of particular interest, because basal activity reduces the chance that the functional impact of mutation was through unintended collateral effects upon folding or structure. In addition to the R280K mutant presented in this dissertation, our group has previously characterized kinetically and/or structurally ten partially active mutants resulting from perturbation to three different active site locations.[7,8] If the predicted impact of each mutation upon the anomeric effect is found to be substantial, these mutants will be used first in our assessment of the relationship between rate and the anomeric effect. Of the available mutants, C271A, at the site of the "essential" cysteine, is the most striking. Its role in electrostatic modulation of a thiolate had not been anticipated [46]. Much of the wild-type activity can be restored by the addition of chloride which binds at the position of the sulfur (thiolate) that is missing in the C271A mutant. This mutant may be particularly useful, because the presence or absence of chloride allows structurally-neutral modulation of the electrostatics at the

catalytic center, supporting comparisons of the predicted anomeric effect with the observed kinetics.

Part b: The AK TSA structure[52] suggests some possible candidates. Four different arginines coordinate the β-oxygens and oxygens of the γ-phosphate analog (NO_3) are coordinated by two arginines, an asparagine and a magnesium ion. Obvious choices for conservative substitutions are lysine, aspartate and glutamine that would perturb ionic / hydrogen bond interactions. Substitution of alternative counter-ions is another possibility. The investigations will be repeated until, when combining parts a & b of Step 4, there are at least six characterized active site variants from which correlations between activity and the anomeric effect or electrostatics should begin to emerge.

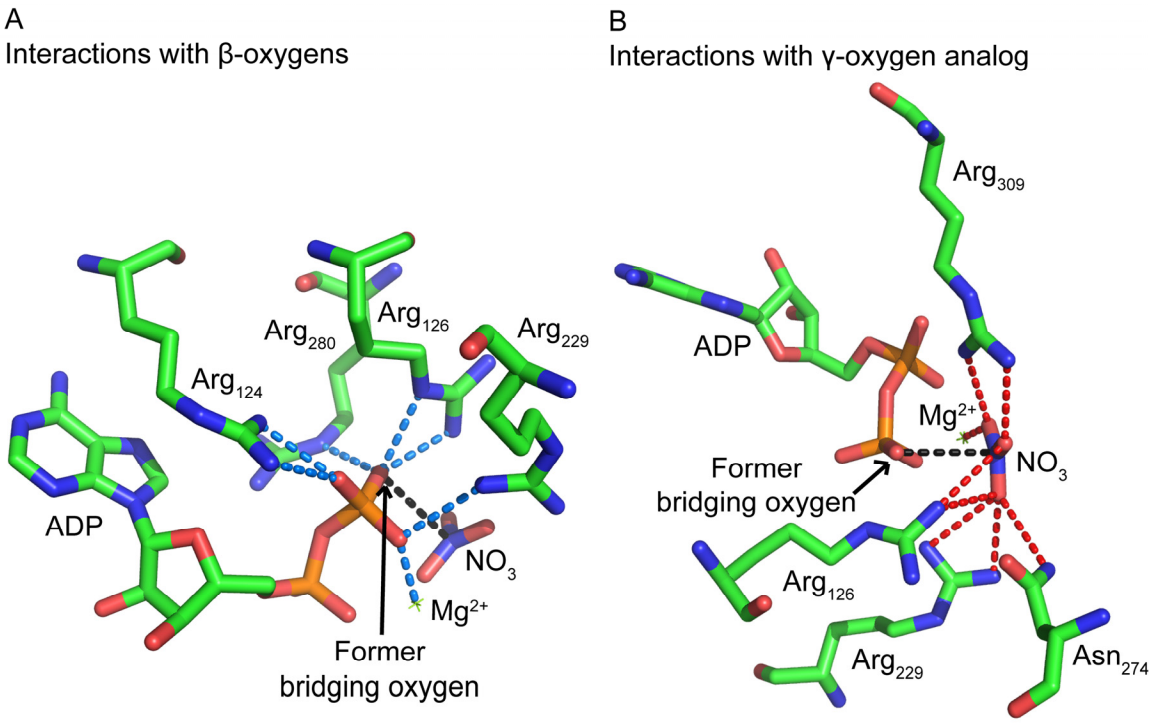


Figure 31. Enzyme-ligand interactions in the transition state analog crystal structure of Arginine kinase (PDB ID 1M15).[52] *Here NO_3 mimicks as the γ-phosphate group in its presumptive transition state position. Dotted black lines indicate path in in-line phosphoryl transfer. In a) interactions between active site residues or Mg2+ and the β-phosphoryl oxygens are shown as blue dotted lines and in b) interactions with γ-oxygens are shown in red. Interactions with β-oxygens enhance the anomeric effect and elongate the scissile O—P bond. In an associative but not dissociative transition-state, interactions at the γ-oxygens would reduce the anomeric effect.*

Potential Outcomes:

There are four potential outcomes; a) both the anomeric effect and electrostatics are correlated with rate, b) neither factor is correlated, c) there is a correlation between rate and electrostatics only and d) there is only a correlation between rate and the anomeric effect.

A correlation for both factors would indicate they may both be important for catalysis. It may not be possible then to differentiate the impact of each factor on catalysis using experiment. However in *in silico* mutagenesis it is often possible to alchemically change individual atoms such that one factor is more specifically affected. Such approaches are discussed in more detail in "Potential challenges".

If neither factor is correlated with rate, it may be because neither factor is important in AK catalysis. Additional mutants may be warranted to cement conclusions as lack of correlation could also result from a limited dataset or mutations that cause collateral damage. Use of *in silico* mutagenesis methods again may help verify results.

A strong correlation with electrostatic energy but none with the anomeric effect would suffice as a negative result, indicating that the anomeric effect is unlikely to play an important role in AK catalysis. Although disproving our original hypothesis such a result would still be quite exciting as this study would still illuminate the origin of arginine kinase's rate enhancement; electrostatics.

A positive correlation with the anomeric effect would support our hypothesis that arginine kinase functions by enhancing the anomeric effect in the ATP substrate to promote catalysis.

Nobody yet has been able to experimentally test multiple aspects of catalysis without running into collateral effects. If either factor is demonstrated unequivocally to be correlated with catalysis, then this work would serve as evidence that catalysis can be dissected into its contributing factors with predictions generated through high level modeling studies followed by thorough experimental validation.

Potential Challenges:

Calculating total energy of the ground state of the reactants is relatively trivial however the transition state calculations called for in the rate calculation in Step 2 can be difficult. The methods available for determining transition states are generally more time consuming, and prone to failure and starting geometries must to be far more accurate than for a standard ground state energy minimum.[139,140] Furthermore, a calculation of rate in the context of a full enzyme using the QM/MM method is a major undertaking where one calculation could take several weeks or even in the more difficult cases, a few months. An easier, but less ideal compromise would be to use ground state calculations of the anomeric effect and rely on experiment for measures of rate.

Step 4, part b carries the risks that mutations will cause a complete loss of activity. To address this, the plan incorporates redundancy in the different amino acid types that may be substituted at each of several locations. For the aim to work, only a proportion has to be amenable to full analysis.

Another issue with our mutagenesis experiments could arise if all potential perturbations equally affect both the anomeric effect and electrostatics. In this case it may be possible to rely instead on *in silico* or "Alchemic" mutagenesis. "Alchemic" atom substitutions are small changes to a single atom or even just a property of an atom that can be performed in a computer model, if not in real life. An example of an Alchemic substitution would be to adjust proton affinity or charge of an atom without changing that atom's size. This method has the added advantage of isolating particular affects without manipulating the overall structure.

This vision further narrows the focus to a single enzyme. This is currently necessary due to the detailed, labor-intensive studies that are now needed. Of course the

real interest is broader, so the hope would be that the studies of AK generate specific hypotheses that can be tested to further establish quantitatively the role of stereoelectronics in the catalysis of a wider array of important enzymes.

The research described here would be challenging and would require multiple years to complete. However, the reward would be significant; this study has the potential to separate out individual factors involved in rate, where computational results are substantiated with considerable integration of experimental data. A positive result would position the anomeric effect as a top contributor to enzyme mechanism and induce a major shift in the existing paradigm of phosphoryl transfer catalysis.

REFERENCES

(1) Wyss, M.; Schulze, A. Health implications of creatine: can oral creatine supplementation protect against neurological and atherosclerotic disease? Neuroscience **2002**, 112, 243.

(2) Vial Creatine Kinase; Nova Science Publishers: New York, 2006.

(3) Stroud, R. M. Balancing ATP in the cell. Nat Struct Biol **1996**, 3, 567.

(4) Murali, N.; Jarori, G. K.; Landy, S. B.; Rao, B. D. Two-dimensional transferred nuclear Overhauser effect spectroscopy (TRNOESY) studies of nucleotide conformations in creatine kinase complexes: effects due to weak nonspecific binding. Biochemistry **1993**, 32, 12941.

(5) Murali, N.; Jarori, G. K.; Rao, B. D. Two-dimensional transferred nuclear Overhauser effect spectroscopy (TRNOESY) studies of nucleotide conformations in arginine kinase complexes. Biochemistry **1994**, 33, 14227.

(6) Zhou, G.; Somasundaram, T.; Blanc, E.; Parthasarathy, G.; Ellington, W. R.; Chapman, M. S. Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. Proc Natl Acad Sci U S A **1998**, 95, 8449.

(7) Pruett, P. S.; Azzi, A.; Clark, S. A.; Yousef, M. S.; Gattis, J. L.; Somasundaram, T.; Ellington, W. R.; Chapman, M. S. The putative catalytic bases have, at most, an accessory role in the mechanism of arginine kinase. J Biol Chem **2003**, 278, 26952.

(8) Gattis, J. L.; Ruben, E.; Fenley, M. O.; Ellington, W. R.; Chapman, M. S. The active site cysteine of arginine kinase: structural and functional analysis of partially active mutants. Biochemistry **2004**, 43, 8680.

(9) Westheimer, F. H. Why nature chose phosphates. Science **1987**, 235, 1173.

(10) McElroy, W. D.; Glass, H. B. Phosphorus Metabolism; Johns Hopkins Press: Baltimore, 1951; Vol. 1.

(11) Miller, D. L.; Westheimer, F. H. The hydrolysis of gamma-phenylpropyl di- and triphosphates. J Am Chem Soc **1966**, 88, 1507.

(12) Voet, D.; Voet, J. G. Biochemistry, 3rd Edition ed.; Wiley, 2003.

(13) Kamerlin, S. C.; Sharma, P. K.; Prasad, R. B.; Warshel, A. Why nature really chose phosphate. Q. Rev. Biophys. **2013**, 46, 1.

(14) Ruben, E. A.; Chapman, M. S.; Evanseck, J. D. Generalized anomeric interpretation of the "high-energy" N-P bond in N-methyl-N'-phosphorylguanidine: importance of reinforcing stereoelectronic effects in "high-energy" phosphoester bonds. J Am Chem Soc **2005**, 127, 17789.

(15) Ruben, E. A.; Plumley, J. A.; Chapman, M. S.; Evanseck, J. D. Anomeric effect in "high energy" phosphate bonds. Selective destabilization of the scissile bond and modulation of the exothermicity of hydrolysis. J. Am. Chem. Soc. **2008**, 130, 3349.

(16) Kirby, A. J. Stereoelectronic Effects; Oxford University Press: Oxford, 1996.

(17) Juaristi, E.; Cuevas, G. The Anomeric Effect (New Directions in Organic and Biological Chemistry). CRC Press, 1995.

(18) Edward, J. Stability of glycosides to acid hydrolysis. Chem. Ind. **1955**, 36, 1102.

(19) McNaught, A. D.; Wilkinson, A. IUPAC Compendium of Chemical Terminology; International Union of Pure and Applied Chemistry: Research Triangle Park, US (NY), 1997.

(20) Gorenstein, D. G.; Findlay, J. B.; Luxon, B. A.; Kar, D. Stereoelectronic control in carbon-oxygen and phosphorus-oxygen bond breaking processes. Ab initio calculations and speculations on the mechanism of action of ribonuclease A, staphylococcal nuclease, and lysozyme. J. Am. Chem. Soc. **1977**, 99, 3473.

(21) Gorenstein, D. G.; Luxon, B. A.; Findlay, J. B. The torsional potential for phosphate diesters. The effect of geometry optimization in CNDO and ab initio molecular orbital calculations. Biochim. Biophys. Acta **1977**, 475, 184.

(22) Gorenstein, D. G.; Taira, K. Stereoelectronic control in peptide bond formation. Ab initio calculations and speculations on the mechanism of action of serine proteases. Biophys. J. **1984**, 46, 749.

(23) Cheng, H.; Nikolic-Hughes, I.; Wang, J. H.; Deng, H.; O'Brien, P. J.; Wu, L.; Zhang, Z. Y.; Herschlag, D.; Callender, R. Environmental effects on phosphoryl group bonding probed by vibrational spectroscopy: implications for understanding phosphoryl transfer and enzymatic catalysis. J Am Chem Soc **2002**, 124, 11295.

(24) Adams, J. A. Kinetic and catalytic mechanisms of protein kinases. Chem. Rev. **2001**, 101, 2271.

(25) Luo, D.; Xu, T.; Watson, R. P.; Scherer-Becker, D.; Sampath, A.; Jahnke, W.; Yeong, S. S.; Wang, C. H.; Lim, S. P.; Strongin, A.; Vasudevan, S. G.; Lescar, J. Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. EMBO J. **2008**, 27, 3209.

(26) Marco-Marin, C.; Gil-Ortiz, F.; Rubio, V. The crystal structure of Pyrococcus furiosus UMP kinase provides insight into catalysis and regulation in microbial pyrimidine nucleotide biosynthesis. J. Mol. Biol. **2005**, 352, 438.

(27) Matte, A.; Tari, L. W.; Delbaere, L. T. How do kinases transfer phosphoryl groups? Structure **1998**, 6, 413.

(28) Parke, C. L.; Wojcik, E. J.; Kim, S.; Worthylake, D. K. ATP hydrolysis in Eg5 kinesin involves a catalytic two-water mechanism. J. Biol. Chem. **2010**, 285, 5859.

(29) Zhu, Y.; Huang, W.; Lee, S. S.; Xu, W. Crystal structure of a polyphosphate kinase and its implications for polyphosphate synthesis. EMBO Rep **2005**, 6, 681.

(30) Ban, C.; Junop, M.; Yang, W. Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. Cell **1999**, 97, 85.

(31) Bi, X.; Corpina, R. A.; Goldberg, J. Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. Nature **2002**, 419, 271.

(32) Lammens, K.; Bemeleit, D. J.; Mockel, C.; Clausing, E.; Schele, A.; Hartung, S.; Schiller, C. B.; Lucas, M.; Angermuller, C.; Soding, J.; Strasser, K.; Hopfner, K. P. The Mre11:Rad50 structure shows an ATP-dependent molecular clamp in DNA double-strand break repair. Cell **2011**, 145, 54.

(33) Schumacher, M. A.; Scott, D. M.; Mathews, II; Ealick, S. E.; Roos, D. S.; Ullman, B.; Brennan, R. G. Crystal structures of Toxoplasma gondii adenosine kinase reveal a novel catalytic mechanism and prodrug binding. J. Mol. Biol. **2000**, 296, 549.

(34) Zhang, Y.; Dougherty, M.; Downs, D. M.; Ealick, S. E. Crystal structure of an aminoimidazole riboside kinase from Salmonella enterica: implications for the evolution of the ribokinase superfamily. Structure **2004**, 12, 1809.

(35) Holden, H. M.; Thoden, J. B.; Timson, D. J.; Reece, R. J. Galactokinase: structure, function and role in type II galactosemia. Cell. Mol. Life Sci. **2004**, 61, 2471.

(36) Qian, X.; Wu, Y.; He, Y.; Luo, Y. Crystal structure of Methanococcus voltae RadA in complex with ADP: hydrolysis-induced conformational change. Biochemistry **2005**, 44, 13753.

(37) Ramon-Maiques, S.; Marina, A.; Gil-Ortiz, F.; Fita, I.; Rubio, V. Structure of acetylglutamate kinase, a key enzyme for arginine biosynthesis and a prototype for the amino acid kinase enzyme family, during catalysis. Structure **2002**, 10, 329.

(38) Gai, D.; Zhao, R.; Li, D.; Finkielstein, C. V.; Chen, X. S. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. Cell **2004**, 119, 47.

(39) Jensen, K. S.; Johansson, E.; Jensen, K. F. Structural and enzymatic investigation of the Sulfolobus solfataricus uridylate kinase shows competitive UTP inhibition and the lack of GTP stimulation. Biochemistry **2007**, 46, 2745.

(40) Kim, S. G.; Cavalier, M.; El-Maghrabi, M. R.; Lee, Y. H. A direct substrate-substrate interaction found in the kinase domain of the bifunctional enzyme, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. J. Mol. Biol. **2007**, 370, 14.

(41) Hur, S.; Bruice, T. C. Enzymes do what is expected (chalcone isomerase versus chorismate mutase). J. Am. Chem. Soc. **2003**, 125, 1472.

(42) Schweins, T.; Geyer, M.; Scheffzek, K.; Warshel, A.; Kalbitzer, H. R.; Wittinghofer, A. Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins. Nat. Struct. Biol. **1995**, 2, 36.

(43) Larsen, T. M.; Benning, M. M.; Rayment, I.; Reed, G. H. Structure of the bis(Mg2+)-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 A resolution: ATP binding over a barrel. Biochemistry **1998**, 37, 6247.

(44) Dittrich, M.; Schulten, K. Zooming in on ATP hydrolysis in F1. J. Bioenerg. Biomembr. **2005**, 37, 441.

(45) Kraft, L.; Sprenger, G. A.; Lindqvist, Y. Conformational changes during the catalytic cycle of gluconate kinase as revealed by X-ray crystallography. J. Mol. Biol. **2002**, 318, 1057.

(46) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. Electrostatic basis for enzyme catalysis. Chem. Rev. **2006**, 106, 3210.

(47) Hasemann, C. A.; Istvan, E. S.; Uyeda, K.; Deisenhofer, J. The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. Structure **1996**, 4, 1017.

(48) Hong, B. S.; Yun, M. K.; Zhang, Y. M.; Chohnan, S.; Rock, C. O.; White, S. W.; Jackowski, S.; Park, H. W.; Leonardi, R. Prokaryotic type II and type III pantothenate kinases: The same monomer fold creates dimers with distinct catalytic properties. Structure **2006**, 14, 1251.

(49) Delbaere, L. T.; Sudom, A. M.; Prasad, L.; Leduc, Y.; Goldie, H. Structure/function studies of phosphoryl transfer by phosphoenolpyruvate carboxykinase. Biochim. Biophys. Acta **2004**, 1697, 271.

(50) Tari, L. W.; Matte, A.; Goldie, H.; Delbaere, L. T. Mg(2+)-Mn2+ clusters in enzyme-catalyzed phosphoryl-transfer reactions. Nat. Struct. Biol. **1997**, 4, 990.

(51) Ruben, E. Stereoelectronic effects in phosphates, FLORIDA STATE UNIVERSITY, 2007.

(52) Yousef, M. S.; Fabiola, F.; Gattis, J. L.; Somasundaram, T.; Chapman, M. S. Refinement of the arginine kinase transition-state analogue complex at 1.2 A resolution: mechanistic insights. Acta Crystallogr D Biol Crystallogr **2002**, 58, 2009.

(53) Berg, J.; Tymoczko, J. L.; Stryer, L. Biochemistry, 5th ed.; W.H. Freeman & Co.: New York, 2002.

(54) Siebold, C.; Arnold, I.; Garcia-Alles, L. F.; Baumann, U.; Erni, B. Crystal structure of the Citrobacter freundii dihydroxyacetone kinase reveals an eight-stranded alpha-helical barrel ATP-binding domain. J. Biol. Chem. **2003**, 278, 48236.

(55) Srivastava, S. K.; Rajasree, K.; Gopal, B. Conformational basis for substrate recognition and regulation of catalytic activity in Staphylococcus aureus nucleoside di-phosphate kinase. Biochim. Biophys. Acta **2011**, 1814, 1349.

(56) Liu, M.; Dobson, B.; Glicksman, M. A.; Yue, Z.; Stein, R. L. Kinetic mechanistic studies of wild-type leucine-rich repeat kinase 2: characterization of the kinase and GTPase activities. Biochemistry **2010**, 49, 2008.

(57) Liu, M.; Girma, E.; Glicksman, M. A.; Stein, R. L. Kinetic mechanistic studies of Cdk5/p25-catalyzed H1P phosphorylation: metal effect and solvent kinetic isotope effect. Biochemistry **2010**, 49, 4921.

(58) Schneck, J. L.; Briand, J.; Chen, S.; Lehr, R.; McDevitt, P.; Zhao, B.; Smallwood, A.; Concha, N.; Oza, K.; Kirkpatrick, R.; Yan, K.; Villa, J. P.; Meek, T. D.;

Thrall, S. H. Kinetic mechanism and rate-limiting steps of focal adhesion kinase-1. Biochemistry **2010**, 49, 7151.

(59) Ullrich, S. J.; Hellmich, U. A.; Ullrich, S.; Glaubitz, C. Interfacial enzyme kinetics of a membrane bound kinase analyzed by real-time MAS-NMR. Nat Chem Biol **2011**, 7, 263.

(60) Olson, A. L.; Cai, S.; Herdendorf, T. J.; Miziorko, H. M.; Sem, D. S. NMR dynamics investigation of ligand-induced changes of main and side-chain arginine N-H's in human phosphomevalonate kinase. J. Am. Chem. Soc. **2010**, 132, 2102.

(61) Stockbridge, R. B.; Wolfenden, R. The intrinsic reactivity of ATP and the catalytic proficiencies of kinases acting on glucose, N-acetylgalactosamine, and homoserine: a thermodynamic analysis. J. Biol. Chem. **2009**, 284, 22747.

(62) Szarek, P.; Dyguda-Kazimierowicz, E.; Tachibana, A.; Sokalski, W. A. Physical nature of intermolecular interactions within cAMP-dependent protein kinase active site: differential transition state stabilization in phosphoryl transfer reaction. J. Phys. Chem. B **2008**, 112, 11819.

(63) Valiev, M.; Yang, J.; Adams, J. A.; Taylor, S. S.; Weare, J. H. Phosphorylation reaction in cAPK protein kinase-free energy quantum mechanical/molecular mechanics simulations. J. Phys. Chem. B **2007**, 111, 13455.

(64) Weinreb, V.; Li, L.; Campbell, C. L.; Kaguni, L. S.; Carter, C. W., Jr. Mg2+-assisted catalysis by B. stearothermophilus TrpRS is promoted by allosteric effects. Structure **2009**, 17, 952.

(65) Dittrich, M.; Hayashi, S.; Schulten, K. On the mechanism of ATP hydrolysis in F1-ATPase. Biophys. J. **2003**, 85, 2253.

(66) Knowles, J. R. Enzyme-catalyzed phosphoryl transfer reactions. Annu. Rev. Biochem. **1980**, 49, 877.

(67) Hill, T. L.; Morales, M. F. Sources of the high energy content in energy-rich phosphates. Arch. Biochem. **1950**, 29, 450.

(68) Wheland, G. W. The Theory of Resonance; John Wiley and Sons, Inc.: New York, 1944.

(69) Hill, T. L.; Morales, M. F. On "High Energy Phosphate Bonds" of Biochemical Interest. J. Am. Chem. Soc. **1951**, 73, 1656.

(70) George, P.; Witonsky, R. J.; Trachtman, M.; Wu, C.; Dorwart, W.; Richman, L.; Richman, W.; Shurayh, F.; Lentz, B. Squiggle-water. An enquiry into the importance of solvation effects in phosphate ester and anhydride reactions. Biochimica et Biophysica Acta **1970**, 223, 1.

(71) Grein, F. The anomeric effect in oxo acids and esters of phosphorus. Journal of Molecular Structure: THEOCHEM **2001**, 536, 87.

(72) Fleming, I. Molecular Orbitals and Organic Chemical Reactions; John Wiley & Sons: Chichester, West Sussex, 2010.

(73) Alabugin, I. V.; Gilmore, K. M.; Peterson, P. W. Hyperconjugation. Wiley Interdisciplinary Reviews: Computational Molecular Science **2011**, 1, 109.

(74) Anet, F. A. L.; Yavari, I. Generalized anomeric effect and barrier to interanl rotation about the oxygen-methylene bond in chloromethyl methyl ether. J. Am. Chem. Soc. **1977**, 99, 6752.

(75) Mo, Y. Computational evidence that hyperconjugative interactions are not responsible for the anomeric effect. Nat Chem, 2, 666.

(76) Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. A Universal Organic Solvation Model. J. Org. Chem. **1996**, 61, 8720.

(77) Kirschner, K. N.; Woods, R. J. Solvent interactions determine carbohydrate conformation. Proc Natl Acad Sci U S A **2001**, 98, 10541.

(78) Kysel, O.; March, P. Influence of polar medium on the anomeric effect.: Quantum chemical study using the ab initio continuum model. Journal of Molecular Structure: THEOCHEM **1991**, 227, 285.

(79) Frushicheva, M. P.; Cao, J.; Chu, Z. T.; Warshel, A. Exploring challenges in rational enzyme design by simulating the catalysis in artificial kemp eliminase. Proc Natl Acad Sci U S A **2010**, 107, 16869.

(80) Florián, J.; Štrajbl, M.; Warshel, A. Conformational Flexibility of Phosphate, Phosphonate, and Phosphorothioate Methyl Esters in Aqueous Solution. J. Am. Chem. Soc. **1998**, 120, 7959.

(81) Yamane, T.; Okamura, H.; Ikeguchi, M.; Nishimura, Y.; Kidera, A. Water-mediated interactions between DNA and PhoB DNA-binding/transactivation domain: NMR-restrained molecular dynamics in explicit water environment. Proteins **2008**, 71, 1970.

(82) Yang, Y.; Cui, Q. The hydrolysis activity of adenosine triphosphate in myosin: a theoretical analysis of anomeric effects and the nature of the transition state. J. Phys. Chem. A **2009**, 113, 12439.

(83) Yang, Y.; Cui, Q. The hydrolysis activity of adenosine triphosphate in myosin: a theoretical analysis of anomeric effects and the nature of the transition state. J Phys Chem A **2009**, 113, 12439.

(84) Rudbeck, M. E.; Nilsson Lill, S. O.; Barth, A. Influence of the molecular environment on phosphorylated amino Acid models: a density functional theory study. J. Phys. Chem. B **2012**, 116, 2751.

(85) Iche-Tarrat, N.; Barthelat, J. C.; Vigroux, A. Theoretical study of specific hydrogen-bonding effects on the bridging P-OR bond strength of phosphate monoester dianions. J. Phys. Chem. B **2008**, 112, 3217.

(86) Petkso, G. A.; Ringe, D. Protein structure and function; New Science Press Ltd.: London, 2004.

(87) Bottoms, C. A.; White, T. A.; Tanner, J. J. Exploring structurally conserved solvent sites in protein families. Proteins **2006**, 64, 404.

(88) Freer, S. T.; Kraut, J.; Robertus, J. D.; Wright, H. T.; Xuong, N. H. Chymotrypsinogen: 2.5-angstrom crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation. Biochemistry **1970**, 9, 1997.

(89) Pedersen, A. K.; Peters, G. G.; Moller, K. B.; Iversen, L. F.; Kastrup, J. S. Water-molecule network and active-site flexibility of apo protein tyrosine phosphatase 1B. Acta Crystallogr. D Biol. Crystallogr. **2004**, 60, 1527.

(90) Phillips, R. S. How does active site water affect enxymatic stereorecognition? Journal of Molecular Catalysis B: Enzymatic **2002**, 19-21, 103.

(91) Jones, P. G.; Kirby, A. J. Simple correlation between bond length and reactivity. Combined use of crystallographic and kinetic data to explore a reaction coordinate. J. Am. Chem. Soc. **1984**, 106, 6207.

(92) Jones, P. G.; Sheldrick, G. M.; Kirby, A. J.; Abell, K. W. Y. Phosphate ester dianions. I. Bis(cyclohexylammonium) propargyl phosphate dihydrate,

2C6H14N+.C3H3O4P2-.2H2O Acta Crystallographica, Section C: Crystal Structure Communications **1984**, C40, 547.

(93) Shurki, A.; Strajbl, M.; Schutz, C. N.; Warshel, A. Electrostatic basis for bioenergetics. Methods Enzymol **2004**, 380, 52.

(94) Strajbl, M.; Shurki, A.; Warshel, A. Converting conformational changes to electrostatic energy in molecular motors: The energetics of ATP synthase. Proc Natl Acad Sci U S A **2003**, 100, 14834.

(95) 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.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision A.02; Gaussian, Inc.: Wallingford CT, 2009.

(96) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for Anion Calculations III. The 3-21+G set for first-row elements, Li-F. J. Comp. Chem. **1983**, 4, 294.

(97) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. Self-Consistant Molecular Orbital Methods. 23. A Polarization Basis Set for Second Row Elements. J. Chem. Phys. **1982**, 77, 3654.

(98) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B Condens. Matter **1988**, 37, 785.

(99) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. Chem. Phys. Lett. **1988**, 153, 503.

(100) Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of ab initio Molecular Potentials for the Prevision of Solvent Effects. Chem. Phys. **1981**, 55, 117.

(101) Kamerlin, S. C.; Florian, J.; Warshel, A. Associative versus dissociative mechanisms of phosphate monoester hydrolysis: on the interpretation of activation entropies. Chemphyschem **2008**, 9, 1767.

(102) Kamerlin, S. C.; Haranczyk, M.; Warshel, A. Are mixed explicit/implicit solvation models reliable for studying phosphate hydrolysis? A comparative study of continuum, explicit and mixed solvation models. Chemphyschem **2009**, 10, 1125.

(103) 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, 2493.

(104) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F. NBO Version 5.9.; 5.9 ed.; Theoretical Chemistry Institute, University of Wisconsin: Madison, Wisconsin, 2001.

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

(106) Foster, J. P.; Weinhold, F. Natural hybrid orbitals. J. Am. Chem. Soc. **1980**, 102, 7211.

(107) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO 3.0 Program Manual; Theoretical Chemistry Institute and Department of Chemistry: University of Wisconsin-Madison.

(108) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO Version 3.1; 3.1 ed.

(109) Kamerlin, S. C.; Warshel, A. On the energetics of ATP hydrolysis in solution. J. Phys. Chem. B **2009**, 113, 15692.

(110) Jeffrey, G. A.; Saenger, W. Hydrogen Bonding in Biological Structures; Springer-Verlag: Berlin, 1994.

(111) Kirby, A. J. The Anomeric Effect and Related Stereoelectronic Effects at Oxygen; Springer-Verlag: Berlin, 1983.

(112) Ruben, E. A.; Chapman, M. S.; Evanseck, J. D. Hydrogen bonding mediated by key orbital interactions determines hydration enthalpy differences of phosphate water clusters. J Phys Chem A **2007**, 111, 10804.

(113) Iancu, C. V.; Borza, T.; Fromm, H. J.; Honzatko, R. B. IMP, GTP, and 6-phosphoryl-IMP complexes of recombinant mouse muscle adenylosuccinate synthetase. J. Biol. Chem. **2002**, 277, 26779.

(114) Kack, H.; Gibson, K. J.; Lindqvist, Y.; Schneider, G. Snapshot of a phosphorylated substrate intermediate by kinetic crystallography. Proc. Natl. Acad. Sci. U S A **1998**, 95, 5495.

(115) Davulcu, O.; Skalicky, J. J.; Chapman, M. S. Rate-limiting domain and loop motions in arginine kinase. Biochemistry **2011**, 50, 4011.

(116) Hagelauer, U.; Faust, U. The catalytic activity and activation energy of creatine kinase isoenzymes. J. Clin. Chem. Clin. Biochem. **1982**, 20, 633.

(117) Iwanami, K.; Iseno, S.; Uda, K.; Suzuki, T. A novel arginine kinase from the shrimp Neocaridina denticulata: the fourth arginine kinase gene lineage. Gene **2009**, 437, 80.

(118) Johnston, I. A.; Goldspink, G. Thermodynamic activation parameters of fish myofibrillar ATPase enzyme and evolutionary adaptations to temperature. Nature **1975**, 257, 620.

(119) Travers, F.; Bertrand, R.; Roseau, G.; Van Thoai, N. Cryoenzymologic studies on arginine kinase: solvent, temperature and pH effects on the overall reaction. Eur. J. Biochem. **1978**, 88, 523.

(120) Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. Covalent Nature of the Strong Homonuclear Hydrogen Bond. Study of the O-H---O System by Crystal Structure Correlation Methods. J. Am. Chem. Soc. **1994**, 116, 909.

(121) Gilli, G.; Gilli, P. The Nature of the Hydrogen Bond: Outline of a Comprehensive Hydrogen Bond Theory; Oxford University Press: New York, 2009.

(122) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. **1999**, 285, 1735.

(123) Fabiola, F.; Bertram, R.; Korostelev, A.; Chapman, M. S. An improved hydrogen bond potential: impact on medium resolution protein structures. Protein Sci. **2002**, 11, 1415.

(124) Summerton, J. C.; Evanseck, J. D.; Chapman, M. S. Hyperconjugation-mediated solvent effects in phosphoanhydride bonds. J. Phys. Chem. A **2012**, 116, 10209.

(125) Blethen, S. L. Kinetic properties of the arginine kinase isoenzymes of Limulus polyphemus. Arch. Biochem. Biophys. **1972**, 149, 244.

(126) Du, X.; Black, G. E.; Lecchi, P.; Abramson, F. P.; Sprang, S. R. Kinetic isotope effects in Ras-catalyzed GTP hydrolysis: evidence for a loose transition state. Proc. Natl. Acad. Sci. U S A **2004**, 101, 8858.

(127) Cepus, V.; Scheidig, A. J.; Goody, R. S.; Gerwert, K. Time-resolved FTIR studies of the GTPase reaction of H-ras p21 reveal a key role for the beta-phosphate. Biochemistry **1998**, 37, 10263.

(128) Du, X.; Sprang, S. R. Transition state structures and the roles of catalytic residues in GAP-facilitated GTPase of Ras as elucidated by (18)O kinetic isotope effects. Biochemistry **2009**, 48, 4538.

(129) Allin, C.; Ahmadian, M. R.; Wittinghofer, A.; Gerwert, K. Monitoring the GAP catalyzed H-Ras GTPase reaction at atomic resolution in real time. Proc. Natl. Acad. Sci. U S A **2001**, 98, 7754.

(130) Allin, C.; Gerwert, K. Ras catalyzes GTP hydrolysis by shifting negative charges from gamma- to beta-phosphate as revealed by time-resolved FTIR difference spectroscopy. Biochemistry **2001**, 40, 3037.

(131) Maegley, K. A.; Admiraal, S. J.; Herschlag, D. Ras-catalyzed hydrolysis of GTP: a new perspective from model studies. Proc. Natl. Acad. Sci. U S A **1996**, 93, 8160.

(132) Pakhomova, S.; Bartlett, S. G.; Doerner, P. A.; Newcomer, M. E. Structural and biochemical insights into the mechanism of fosfomycin phosphorylation by fosfomycin resistance kinase FomA. Biochemistry, 50, 6909.

(133) Davies, D. R.; Hol, W. G. The power of vanadate in crystallographic investigations of phosphoryl transfer enzymes. FEBS Lett **2004**, 577, 315.

(134) Niu, X.; Bruschweiler-Li, L.; Davulcu, O.; Skalicky, J. J.; Bruschweiler, R.; Chapman, M. S. Arginine kinase: joint crystallographic and NMR RDC analyses link substrate-associated motions to intrinsic flexibility. J Mol Biol **2011**, 405, 479.

(135) Hratchian, H. P.; Schlegel, B. H. Theory and Applications of Computational Chemistry: The First Forty Years; Elsevier B.V., 2005.

(136) Hansen, D. E.; Knowles, J. R. The stereochemical course of the reaction catalyzed by creatine kinase. J Biol Chem **1981**, 256, 5967.

(137) Knowles, J. R. Enzyme-catalyzed phosphoryl transfer reactions. Annu Rev Biochem **1980**, 49, 877.

(138) Glennon, T. M.; Villa, J.; Warshel, A. How does GAP catalyze the GTPase reaction of Ras? A computer simulation study. Biochemistry **2000**, 39, 9641.

(139) Leach, A. R. Molecular Modelling Principles and Applications, 2nd ed.; Pearson Education Limited: Essex, England, 2001.

(140) Peng, C.; Schlegel, B. Combining Synchronous Transit and Quasi-Newton Methods to Find Transition States. Israel Journal ofChemistry **1993**, 33, 449.

APPENDIX I: EFFORT STATEMENT

Chapter 2: Hyperconjugation-mediated solvent effects in phosphoanhydride bonds. This work was originally published in J. Phys Chem. A:

Summerton, J. C.; Evanseck, J. D.; Chapman, M. S. Hyperconjugation-Mediated Solvent Effects in Phosphoanhydride Bonds. J. Phys. Chem. A 2012, 116, 10209.

MSC and JDE were instrumental in editing of the manuscript. All other work was done by JCS.

Chapter 3: Enzymes Promote Nucleotide Phosphoryl Transfer through Interactions with the Bridging Oxygen that Enhance Hyperconjugation.

The original idea to perform a database search to determine whether enzymes are configured for anomeric bond destabilization was from MSC. The design and execution of the database analysis was done by JCS.

The original idea for the experimental rate measurements on the Arginine kinase mutant and design of the mutation were done by JCS. The molecular biology associated with the mutagenesis work was performed by GMM. The protein preparation and kinetics measurements were performed by JCS.

Design of the quantum mechanics calculations and their execution was performed by JCS. The Manuscript was written by JCS and primarily edited by MSC. Edits were also done by JDE.