Florida State University Libraries

2007

Stereoelectronic Effects in Phosphates

Eliza A. Ruben

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

STEREOELECTRONIC EFFECTS IN PHOSPHATES

By

ELIZA A. RUBEN

A Dissertation submitted to the
Program in Molecular Biophysics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Summer Semester, 2007

The members of the Committee approve the Dissertation of Eliza A. Ruben defended on June 11th 2007.

                          _____

Michael S. Chapman
Professor Co-Directing Dissertation

_____

W. Ross Ellington
Professor Co-Directing Dissertation

_____

Robert L. Fulton
Outside Committee Member

_____

Timothy A. Cross
Committee Member

_____

Huan-Xiang Zhou
Committee Member

Approved:

_____
Timothy M. Logan, Interim Chair, Institute of Molecular Biophysics

The Office of Graduate Studies has verified and approved the above named committee members.

To my father

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Molecules containing the phosphate ($O-PO_3^{2-}$) moiety are ubiquitous in biochemistry. Phosphoryl transfer reactions that break and form the O—P phosphoryl bond are central to biological processes as diverse as energy metabolism and signal transduction. As described by Westheimer, the utility of phosphates stem from their ability to be kinetically stable while thermodynamically unstable. This dissertation uses electronic structure theory to investigate, at an elementary chemical level, the thermodynamic and kinetic properties of phosphate esters in an attempt to answer the question, "Why nature chose phosphates?". Chapter 1 formulates the question to be answered. Chapter 2 provides the underlying theoretical background to the computational methods employed. In Chapter 3, the anomeric effect, a stereoelectronic effect is first identified as a contributor to the high-energy status of N-phosphoryl-guanidines using electronic structure methods. In Chapter 4 it is further found that the anomeric effect can contribute to the thermodynamic poise of a range of phosphates. Chapter 5 investigates the connection between phosphoryl transfer mechanisms and the anomeric effect. It is found that the anomeric effect promotes O—P bond cleavage and plays a dominant role in the dissociative mechanism of phosphoryl transfer. The impact of other stereoelectronic effects such as hyperconjugation upon the hydrogen bonding properties of phosphates is also examined in Chapter 6. Compelling evidence is obtained suggesting the role of the O—P bond weakening anomeric effect in discriminating phosphoryl transfer potentials and controlling reaction rates in a range of biologically important phosphoryl compounds. Strong correlations between phosphoryl transfer potentials, rates of reaction in solution, O—P bond weakening, and the magnitude of the $n(O) \rightarrow \sigma^*(O-P)$ anomeric effect is shown. This dissertation articulates a fundamental property of phosphates that may provide an answer to the age old question of "Why nature chose phosphates".

INTRODUCTION

Why Did Nature Choose Phosphates?

Phosphates dominate the natural world with phosphoryl transfer reactions forming the centerpiece of biochemical reactions.[1] The most essential biochemical phosphate is adenosine triphosphate or ATP. The importance of ATP to biology is underscored by Boyer in his 1997 Nobel address who estimated that the synthesis of ATP is the most prevalent chemical reaction in the human body.[2] The importance of phosphates stem from their energy-transducing properties. Energy derived from the hydrolysis of phosphates such as ATP are involved in biological processes as diverse as macromolecule synthesis, muscle contraction, active transport and apoptosis.[3] Other phosphates such as GTP play major roles in signal transduction and phospodiesters form the foundation of genetic material.[3] As elegantly described by Westheimer, the ubiquity of phosphates in nature prompts the question, why did nature choose phosphates and almost no other chemically similar group? [4]

Westheimer pointed out that phosphoryl transfer reactions in nature were essentially nucleophilic displacement reactions requiring good leaving groups.[4] Yet, synthetic chemists almost always use better leaving groups such as chlorides, bromides and iodides.[4] Despite this greater propensity as better leaving groups, nature almost never uses such groups for nucleophilic displacement reactions. Phosphates and their derivatives are almost always used in natural systems. What is it about phosphates that command this place in nature?

Westheimer followed up by proposing that the utility of phosphates stemmed from their ability to be both thermodynamically unstable while kinetically stable.[4] The thermodynamic instability of phosphates is well known as a source of energy for biochemical processes.[5] The kinetic stability or slow reactivity of phosphates allows this reservoir of energy to be stored (without reacting) before being activated by enzymatic environments.[4] Although this idea of kinetic stability vs thermodynamic instability provides a good overall sense of

why phosphates are so useful to nature it does not really answer the question of why nature chose phosphates. There is yet to be articulated a fundamental chemical property of phosphates that nature exploits to achieve this thermodynamic instability and kinetic stability.

Thermodynamic Instability Of Phosphates

The idea of phosphates being a primary biochemical energy provider was first coined by Lipmann in 1941 where he proposed the classical view of high-energy bonds or high-energy phosphates.[5] The high-energy bond was described as a bond whose hydrolysis yielded large exothermic free energies and the energy derived from was used to drive biochemical processes. Only selected phosphates can be described as high-energy as not all phosphates produce large exothermic free energies when hydrolyzed. Examples of high-energy bonds include the N—P bond of arginine phosphate and creatine phosphate and the O—P bond of acetyl phosphate, carbamyl phosphate and the ubiquitous ATP.

The term high-energy bond has been confused as describing its bond energy which is defined as the energy required to break, not hydrolyze, a covalent bond [6]. Since this process is energy-requiring, not energy-releasing, the term high-energy has met with its share of detractors. It is important note that hydrolysis is made up of a bond breaking process that requires energy and a bond forming process that produces energy. A hallmark of high-energy bonds is their lability or weakness. Breaking a high-energy bond requires little energy. So the energy produced by bond-forming process dominates the overall free energy of hydrolysis. In contrast, low-energy phosphate bonds are significantly stronger and a greater energy requirement to break and the overall hydrolysis free energy of low-energy bonds is significantly less exothermic. Thus, as described by Chance et. al., [7] the term high-energy remains a useful phrase to describe a bond whose hydrolysis releases a large amount of energy.

Lipmann proposed the concept of high-energy bonds in accordance with the known experimentation of his time. It was not till later that it was discovered that free energies of hydrolysis were variable. Firstly, it was shown that free energies of hydrolysis also depends on the concentration of products and

reactants as overall free energies of hydrolysis are dependent on the equilibrium constant, equations 1 and 2.[8]

$$ATP^{4-} + H_2O \rightleftharpoons ADP^{3-} + HPO_4^{2-} + H^+ \quad \text{[eqn 1]}$$
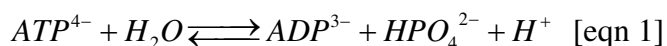
$$\Delta G = \Delta G^\circ + RT \ln \frac{[ADP][HPO_4^{2-}][H^+]}{[ATP^{4-}][H_2O]} \quad \text{[eqn 2]}$$

Enzymes are known for exploiting this capability to change overall free energies by binding select substrates in order to change the equilibrium constant. Secondly, free energies of hydrolysis are also dependent upon the environment with cation coordination, pH changes and enzyme active site environments all imparting differences in hydrolysis free energies of phosphates.[9,10] Deriving the energy contained in these high-energy bonds does not always require hydrolysis. Phosphoryl transfer reactions between high and low-energy phosphates allowed the coupling of energy-releasing processes to energy-requiring processes in order to drive them to completion.[3,9,11] This has led Lipmann's postulate that it was the chemical nature of the high-energy bond that determined the magnitude of the free energies of hydrolysis and provided energy for biological processes to be questioned.[10]

Despite changes afforded by the equilibrium constant, the standard free energy remains the backdrop of overall hydrolysis free energies. Why would nature want to work harder, whether by means of raising concentrations of cellular phosphates or selective binding, in order to provide the necessary energy for biological processes instead of depending, at least partially, on the thermodynamic poise of phosphates alluded to by Lipmann? Moreover the notion that high-energy bonds cannot account for differences in hydrolysis free energies seems to be derived from the belief that the high-energy bond has constant chemical properties and so cannot account for such varying free energies of hydrolysis. What if there is a property of high-energy bonds that could account for such variability?

The high-energy character of relevant phosphates has been attributed to three main factors; resonance effects, solvation and electrostatics. Resonance

factors were first proposed as an explanation for the high-energy character by Kalckar in 1940.[12] Kalckar states that high-energy phosphates displayed a characteristic feature, which he referred to as opposing resonance. Specifically, the -N- and -O- linkages at the bridge of high-energy esters were influenced by opposing resonance between the organic group at one end and the phosphate group at the other. Both the organic group and the phosphate group make demands on the bridging -N- and -O- for their independent resonating systems, Figure 1. In his own words, "Resonance of type A opposes that of type B leading to less resonance energy for the groups combined as an ester than for them as independent."

Type A

Type B
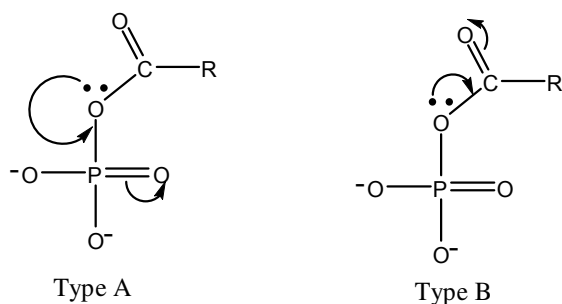
Figure 1.1. A representation of Kalckar's opposing resonance theory.

The increased stability of the hydrolyzed product in comparison to the instability of the -O- and -N- linkage is responsible for the large exothermic free energies of hydrolysis released when such linkages are hydrolyzed. The main weakness of the opposing resonance theory stems from the fact that it is not quantitative and has yet to be proven to explain the thermodynamic poise of a range of phosphates.

George et. al. proposed that solvation factors were more important in imparting high-energy character to phosphates than opposing resonance.[10,13,14] Key to this explanation is the fact that a more solvated molecule is more stable than a less solvated molecule.[10,13,14] George and coworkers found that the solvation energy of high-energy phosphates was much lower than their hydrolysis products.[10,13,14] George and coworkers were also the first to notice the impact of ionic form or protonation state upon free energies of hydrolysis. The more ionized the phosphate, the greater the magnitude of the hydrolysis free

energies. George et. al. saw this as further proof of the impact of solvation upon free energies of hydrolysis. They proposed that the relationship between protonation state and free energies of hydrolysis was related to entropic effects stemming from interactions between ion and solvent, where a greater entropic energy term led to more exothermic free energies of hydrolysis.[10,13,14] While this reason is certainly more quantitative than opposing resonance, solvation factors may be a consequence of inherent properties of phosphates. Solvation has also only been shown to provide reasons for a select range of phosphates.

The final factor often used to explain high-energy character in ATP is electrostatics. Here adjacent negatively charged oxygen atoms repel each other weakening the -O- linkage decreasing the stability of the ATP reactant relative to the hydrolysis product. This reason cannot explain the high-energy character of phosphates not possessing adjacent oxygen atoms such as arginine, creatine, carbamyl and acetyl phosphates.

<u>Kinetic Stability of Phosphates</u>

Equally important as the ability of phosphates to exhibit thermodynamic poise and provide the energy needed to drive biochemical processes is their kinetic stability. The kinetic stability of phosphates refers to the low reactivity of phosphates in solution compared to the rate enhancements attained in enzymatic environments. This property allows phosphates to be stored in solution, transported where needed before being activated for use by enzymes.

Enzymes accelerate biochemical reactions by reducing the activation energy barrier or stabilizing the transition state. The catalytic power of phosphoryl transfer enzymes can reach heights of 10^{21} fold rate accelerations.[15] Enzymes achieve catalysis through a variety of ways such as acid-base catalysis, substrate preorganization, covalent and electrostatic catalysis.[16] Phosphoryl transfer enzymes are no exception and utilize different ways of achieving catalysis.

The structure of the arginine kinase transition state analogue shows substrates aligned to within 3° for optimal phosphoryl transfer.[17] This implicates substrate pre-alignment as an important strategy utilized by bimolecular

phosphoryl transfer enzymes. The haloalkane dehydrogenase (HAD) family of phosphotransferases utilize strategies from simple acid base catalysis to solvent exclusion to catalyze phosphoryl transfer.[1]

Despite the myriad of strategies enzymes employ to catalyze phosphoryl transfer, the original question 'Why did nature use phosphates and not other groups?', is still relevant. The ability of the O—P bond to be responsive to environmental changes may be an important property exploited by enzymes to achieve catalysis. The trapping of the pentacovalent phosphorane as an intermediate in enzymatic phosphoryl transfer showed O—P bond lengths stretched to 2.1 Å.[18] If the O—P bond were stable and invariable, such as a C—C bond in methane, then phosphoryl transfer enzymes would have extra 'work' to do to achieve catalysis.

Even in solution phosphates demonstrate a remarkable variability in reaction rates. Changes in pH alone can induce rate enhancements of 10^3.[19] Organic solvents can increase hydrolysis rates by almost 10^6 fold.[20,21] The chemical nature of phosphates must be able to account for such variability.

Stereoelectronic Effects

Stereoelectronic effects are the chemical and kinetic consequences stemming from interactions between atomic orbitals. Such effects are prominent in organic chemistry where it has been shown that they profoundly affect chemical reactivity and conformation.

One of the more common types of stereoelectronic effect is the anomeric effect. The anomeric effect was first recognized as the preference of polar groups attached to the C1 carbon of galactopyranosyl (sugar) derivatives to attain an axial rather than equatorial position. Later orbital rationalizations for the anomeric effect showed that this preference was due to highest occupied molecular orbital (HOMO) of the polar group the antibonding s-orbital of the C-X bond.

This orbital rationalization was generalized to describe the geometric consequences of any system containing the Lp-X-A-Y moiety where "Lp" represents a lone pair, "X", any heteroatom, "A" an electropositive atom and "Y"

and electronegative atom [22-25]. Electron density from the lone pair is delocalized into the antibonding A—Y orbital creating a stabilizing effect. Critical to the maximization of the anomeric effect is the antiperiplanar arrangement of the lone pair and antibonding orbital as seen in Figure 3. This arrangement maximizes overlap between the orbitals allowing for maximal electron delocalization.

The geometric consequences of this interaction is seen as a lengthening of the A—Y bond; increase in electron density in an antibonding orbital always has a weakening effect upon that bond. Another geometric consequence of the anomeric effect is the shortening of the X—A bond. The classical consequence of the generalized anomeric effect is the double bond/ no bond resonance structure as shown in Figure 3.
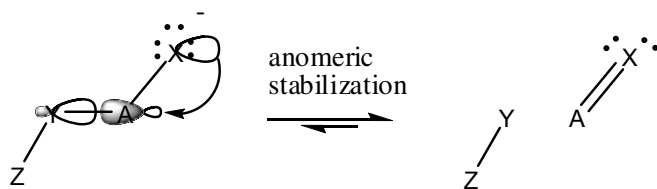
Figure 1.2. The generalized anomeric effect.

Phosphoryl transfer reactions take place through three distinct mechanisms. The dissociative mechanism via a metaphosphate intermediate, a fully associative mechanism requiring no intermediate and a concerted mechanism that proceeds via a phosphorane intermediate, Figure 1.3.

The dissociative mechanism, in particular, closely resembles the path of the anomeric effect in phosphates, since one extra double bond is formed in the phosphoryl moiety and the O—P bond removed. Can stereoelectronic effects, such as the anomeric effect, explain the variability in thermodynamic and kinetic parameters of phosphates and answer the age old question of why nature chose phosphates?

The possibility that such stereoelectronic factors could be a factor in O—P bond cleavage and subsequently reaction rate enhancements was first proposed by Gorenstein et. al. in 1977 when investigating the reactions of phosphate diesters.[26-28] Here it was first noted the selective weakening O—P ester bonds anti-periplanar to lone pairs on directly bonded oxygen atoms when it was found

that O—P bond lengths changed along a stereoelectronically controlled reaction path. Further theoretical and experimental studies on the conformational preference of phosphates, such as in nucleic acids and nucleic acid counterparts, also suggested a stereoelectronic contribution. [29-32] However, the studies of that time lacked the methodology to actually quantify stereoelectronic effects in terms of lone pair to antibond delocalization. The best one could do was identify low-energy conformers consistent with the arrangement of antiperiplanar lone pairs to the ester bond and suggest stereoelectronic effects as the source. Perhaps due to this lack of quantitative information, rigorous studies connecting stereoelectronic effects (in terms on lone pair to antibond delocalization) to characteristic thermodynamic, kinetic and conformational behaviour [33] of important biological phosphates have been rare. As mentioned by Westheimer [4], the importance of phosphates to nature stem from their kinetic stability and thermodynamic instability. Stereoelectronic effects in the context of n → σ*(A—Y) delocalizations are largely absent in the current scientific discourse on both issues.[1,34-41]

The natural bond orbital analysis (NBO) method [42] allows quantification of stereoelectronic effects in strict orbital to orbital delocalization terms. In NBO terminology, the anomeric effect is quantified as an n → σ*(A—Y) interaction with donation of electron density into the antibonding A—Y orbital leading to bond weakening effects. The O—P—O moiety along the phosphoryl compound ideally satisfies all conditions needed for the anomeric effect to occur.[43] The O—P bond is polar and each oxygen atom on the phosphoryl group has up to three lone pairs available for the donation of electron density into the antibonding σ*(O—P) orbital.

Here we use the NBO method to revisit old problems such as the *high energy* nature and reactivity of phosphates.

Dissociative mechanism

metaphosphate intermediate

Associative mechanism

associative transition state
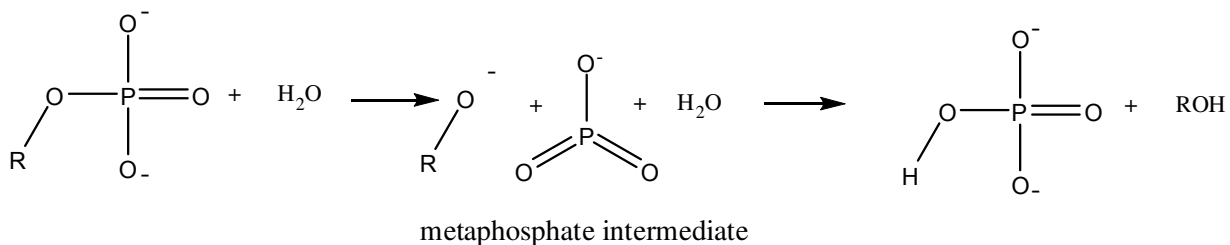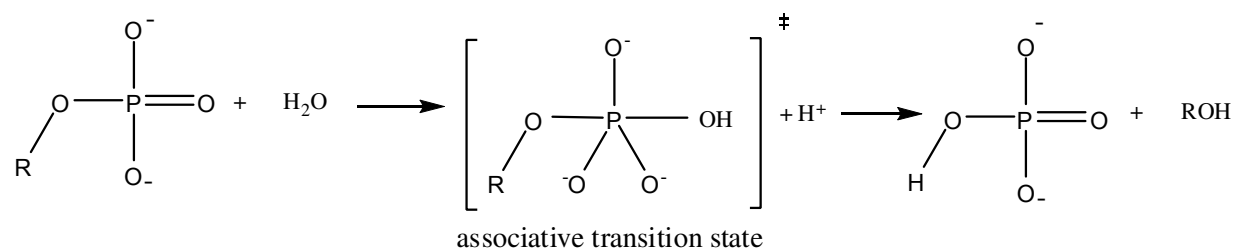
Concerted via a phosphorane intermediate

phosphorane intermediate
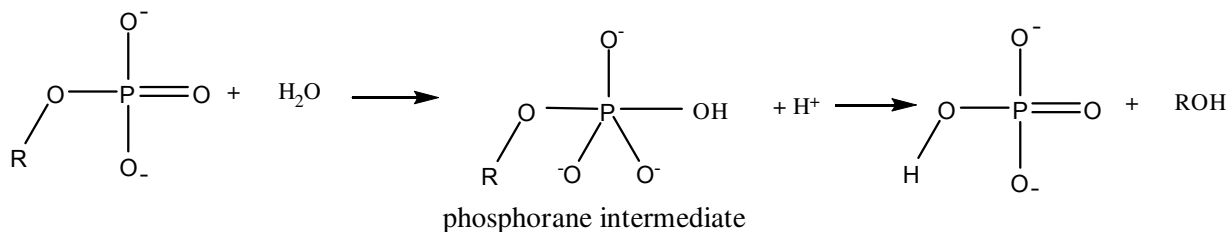
Figure 1.3. Phosphoryl transfer mechanisms

We first show that stereoelectronic effects can account for the *high energy* nature of N-phosphoryl-guanidines. This work is extended to a series of phosphates, followed by an assessment of phosphoryl transfer mechanisms in the context of the anomeric effect. A study of hydrogen bonding in phosphates in the context of orbital overlap is also examined.

COMPUTATIONAL METHODS

   Ab initio quantum mechanical methods were used to examine the electronic structure of phosphates from a theoretical viewpoint. Natural bond orbital (NBO) methods were used to identify and quantify stereoelectronic effects associated with kinetic and thermodynamic parameters of phosphates.

<u>Quantum Chemical Methods</u>

   The primary computational method used in this work is quantum mechanics. Quantum mechanics uses the Schrodinger equation, equation 3, to calculate the electronic structure of a molecule.

$$\hat{H}\psi = E\psi \quad \text{[eqn 3]}$$

where $\hat{H}$ represents the Hamiltonian operator, ψ represents the wavefunction and E the energy of the system.

   The Hamiltonian operator represents the kinetic and potential energy contributions of the electrons and nuclei of a system. For an N electron system, the Hamiltonian takes the form of,

$$\hat{H} = -\frac{\hbar^2}{2}\sum_{\alpha}\frac{1}{m_{\alpha}}\nabla_{\alpha}^2 - \frac{\hbar^2}{2m_e}\sum_{i}\frac{1}{m_i}\nabla_{i}^2 + \sum_{\alpha}\sum_{\beta>\alpha}\frac{Z_{\alpha}Z_{\beta}e'^2}{r_{\alpha\beta}} - \sum_{\alpha}\sum_{i}\frac{Z_{\alpha}e'^2}{r_{i\alpha}} + \sum_{j}\sum_{i>j}\frac{e'^2}{r_{ij}} \quad \text{[eqn 4]}$$

where α and β refer to nuclei and i and j refer to electrons. And where,

$$-\frac{\hbar^2}{2}\sum_{\alpha}\frac{1}{m_{\alpha}}\nabla_{\alpha}^2$$ refers to the kinetic energy of the nuclei

$$-\frac{\hbar^2}{2m_e}\sum_{i}\frac{1}{m_i}\nabla_{i}^2$$ refers to the kinetic energy of the electrons

$$\sum_{\alpha}\sum_{\beta>\alpha}\frac{Z_{\alpha}Z_{\beta}e'^2}{r_{\alpha\beta}}$$ refers to the potential energy due to repulsions between nuclei α and β

$$-\sum_{\alpha}\sum_{i}\frac{Z_{\alpha}e'^2}{r_{i\alpha}}$$ refers to the potential energy due to attractions between nuclei and electrons

$\displaystyle\sum_{j}\sum_{i>j}\frac{e'^2}{r_{ij}}$ refers to the potential energy due to repulsions between electrons

In order to make the solution of the Schrodinger equation more tractable, the Born-Oppenheimer approximation is used. Here, because the electrons move at a speed much faster than the nuclei, the nuclei are considered fixed and the Hamiltonian reduces to

$$\hat{H} = -\frac{\hbar^2}{2m_e}\sum_{i}\frac{1}{m_i}\nabla_i^2 - \sum_{\alpha}\sum_{i}\frac{Z_\alpha e'^2}{r_{i\alpha}} + \sum_{j}\sum_{i>j}\frac{e'^2}{r_{ij}} \quad \text{[eqn 5]}$$

Solving the Schrodinger equation analytically is still impossible for polyatomic systems and further approximations in the form of the variation and perturbation theory are usually employed. The variational theorem states that the energy of an exact wavefunction is always less than or equal to the energy calculated by an approximate wavefunction. By choosing an appropriate wavefunction with flexible parameters and minimizing those parameters a less expensive solution to the Schrodinger equation can be sought. In the perturbation approximation, a small correction called a perturbation is introduced into the Hamiltonian. By calculating the corrections to the wavefunction and energy at different orders an approximate solution to the Hamiltonian is obtained.

Hartree-Fock Method

One of the earliest and most common electronic structure methods of solving the Schrodinger equation is the Hartree-Fock method. The basis of the Hartree-Fock method was initiated by Hartree in 1828 when he proposed a self-consistent field method to describe the movement of one electron with regard to all other electrons and nuclei in a polyatomic system. Key to the Hartree-Fock method is the description of an N electron wavefunction as a Hartree product.

$$\psi(x_1, x_2, \ldots) = \theta_1(x_1)\theta_2(x_2)\ldots \quad \text{[eqn 6]}$$

where x_1 represents Cartesian coordinates and θ_1 represents electron orbitals. To account for the antisymmetry requirement as demanded by Pauli's exclusion principle, John Slater introduced a version of this representation called a Slater

determinant, equation 7. The Slater determinant successfully antisymmetrizes the wavefunction.

$$\psi(x_1, x_2, \ldots) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \theta_1(x_1), \theta_2(x_1), \ldots\ldots\ldots\ldots\theta_n(x_1) \\ \theta_1(x_2), \theta_2(x_2)\ldots\ldots\ldots\ldots\theta_n(x_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \theta_1(x_n), \theta_2(x_n)\ldots\ldots\ldots\ldots\theta_n(x_n) \end{vmatrix} \quad \text{[eqn 7]}$$

The Hamiltonian is rewritten in terms of one and two electron integrals to accommodate the Slater determinant.

$$\hat{F} = -\frac{\hbar^2}{2m_e} \sum_i \frac{1}{m_i} \nabla^2 - \sum_\alpha \sum_i \frac{Z_\alpha e'^2}{r_{i\alpha}} + j + \sum_{j=1}^{n/2} 2\hat{J}(1) - \hat{K}_j(1) \quad \text{[eqn 8]}$$

$\hat{J}$ refers to the Coulomb operator accounting for potential energies of interaction and $\hat{K}_j$ refers to the exchange operator responsible for exchange interaction between 2 electrons.

The limitation of the Hartree-Fock method is its failure to account for correlation between opposite electron spins. As such the Hartree-Fock method is described to lack electron correlation. The difference between the exact energy of a system and the Hartree-Fock energy is therefore known as the correlation energy.

$$E_{corr} = E_{exact} - E_{HF} \quad \text{[eqn 9]}$$

Møller-Plesset Theory

Møller-Plesset theory uses the perturbation method to account for electron correlation. A perturbation representing the difference between the true interelectronic interactions and that of the Hartree-Fock interelectronic representation is introduced into the Fock operator. The zeroth and first order terms represent the Hartree-Fock energy. The second order term takes the form of equation 10.

$$E_0^{(2)} = \sum_{s \neq 0} \frac{\left|\left\langle \psi_s^{(0)} \left| \hat{H} \right| \phi_0 \right\rangle\right|^2}{E_0^{(0)} - E_s^{(0)}} \quad \text{[eqn 10]}$$

$\psi_s^{(0)}$ are all possible Slater determinants. The wavefunction ϕ_0 is a Hartree-Fock wavefunction that includes both occupied and virtual (unoccupied) spin orbitals. The wavefunction ϕ_0 provides an avenue for an occupied spin orbital to be replaced by an unoccupied spin orbital. In this way electron spins can be correlated and the second order correction provides an energetic consequence of this correlation.

<u>Density Functional Theory</u>

Density functional theory is based on the premise that all ground state descriptors such as energy, wavefunction and electronic structure are functions of electron density and that this electron density in turn is a function of just three variables, equation 10.

$$E(\rho) = \bar{T}(\rho) + \bar{V}_{Ne}(\rho) + \bar{V}_{ee}(\rho) \quad \text{[eqn 10]}$$

where $\bar{T}(\rho)$ is the kinetic energy, $\bar{V}_{Ne}(\rho)$ the nuclear-electron interaction and the $\bar{V}_{ee}(\rho)$ electron-electron interaction. A term that includes exchange and correlation is also included making the general form,

$$E(\rho) = \bar{T}(\rho) + \bar{V}_{Ne}(\rho) + \bar{V}_{ee}(\rho) + E_{EX}(\rho) \quad \text{[eqn 11]}$$

A significant advantage of density functional theory is that because electrons are accounted for as density instead of individual electron populations, the method is much less computationally demanding than true correlated methods.

Basis Sets

A basis set is simply a collection of functions used to represent atomic orbitals. A linear combination of these atomic orbitals then represents a wavefunction of a polyatomic system, equation 12.

$$\psi_i = \sum_{\mu=1}^{N} c_{\mu i} \phi_\mu \quad \text{[eqn 12]}$$

Where $c_{\mu i}$ are molecular orbital expansion coefficients and ϕ_μ are the basis functions.

The basis function is essentially a mathematical description of the space an electron is confined to. Increasing the number of basis functions per atom results in a larger basis set that more accurately describes the molecular orbital in space since they provide greater flexibility in the space that electrons can occupy. The most common type of basis functions include Slater type functions which use exponential terms to approximate the radial function of electrons, and Gaussian functions.

In addition to size considerations, shape considerations also influence the utility of basis sets in accurately describing the electronic structure of molecules. To allow orbitals to change shape, polarization functions that add angular momentum to orbitals can be used. Anions are notorious for having electron density far away from nuclei. To treat this problem, diffuse functions that can account for disperse electron density is used.

Natural Bond Orbital Analysis

Natural bond orbital analysis is a technique for studying molecular structure, in particular hybridization and delocalization effects. It is based on the concept of describing molecular structure in terms of electrons localized onto bonds and orbital sans as such provides a more intuitive picture of electronic structure. In NBO analysis the chosen basis set or wavefunction is transformed into a series of localized basis sets such as natural atomic orbitals (NAO),

natural hybrid orbitals (NHO), natural bond orbitals (NBO) and natural localized molecular orbitals (NLMO). The chosen basis set is transformed via orthogonalization and normalization into a complete orthonormal basis set of orbitals that can be expanded according to equation 13.

$$\gamma(\vec{r} \mid \vec{r}') = \sum_{i,j} \gamma_{ij} \theta_i(\vec{r}) \theta_j *(\vec{r}') \text{ [eqn 13]}$$

where $\gamma(\vec{r} \mid \vec{r}')$ represents the density matrix. Then as proposed by Lowdin, the operator $\hat{\gamma}$ can be described by equation 14.

$$\hat{\gamma} \theta_i = n_i \theta_i \text{ [eqn 14]}$$

where n_i represents the occupancy of the eigenorbital θ_i. This equation describes the density matrix as the sole function of the occupancy of the natural orbital θ_i. A set of natural atomic orbitals is obtained by finding the θ_i with maximum occupancy i.e. ~ 2 in the density matrix $\hat{\gamma}$ of an atom. Similarly a set of natural bond orbitals is obtained by searching a density matrix $\hat{\gamma}$ of a bonding region. Because NBOs and NAOs of the highest occupancy are obtained by this method, a Lewis structure representation of lone pairs and bonds is obtained. Each NBO can be further decomposed into natural hybrid orbitals of the form

$$\sigma_{AB} = c_A h_A + c_B h_B \text{ [eqn 15]}$$

where h_A and h_B are the hybrid orbitals and c_A and c_B are the polarization coefficients. An artifact of this decomposition is out of phase non-Lewis orbitals in the form of

$$\sigma^*_{AB} = c_B h_A - c_A h_B \text{ [eqn 16]}$$

These non-Lewis NBOs represent empty valence shells and are termed antibonds. In Lewis structure representation such orbitals are normally unoccupied, but departures from this idealized picture occur leading to small

occupancies of these antibonds that generally amount to < 1%. In effect these small occupancies are electron delocalizations. The utility of these antibonding orbitals is that because their occupancy is minimal, they can be treated as a perturbation using perturbation theory. This then allows the delocalization effects to be quantified by second order perturbative expressions.

$$\Delta E^{(2)} = -2 \frac{\left\langle \sigma \left| \hat{F} \right| \sigma^* \right\rangle^2}{\varepsilon_{\sigma^*} - \varepsilon_{\sigma}} \quad \text{[eqn 17]}$$

In this NBO allows the quantification of delocalization effects between individual orbitals that other treatments of delocalization or charge transfer do not allow.

The other most commonly applied method of studying charge transfer interactions is the Kitaura-Morokuma method.[44] The main criticism of the NBO method is that it has a tendency to overestimate charge transfer effects. A classic test case of this is in the study of charge transfer effects in the water dimmer. The charge transfer given by NBO is given as -9.6 kcal/mol whilst that given by the Kitaura-Morokuma method is -1.8 kcal/mol using the same 6-31G* basis set.

The origin of this discrepancy lies in the way charge transfer is treated between the two methods. In order to define charge-transfer the total atomic orbital space is divided into mutually orthogonal subspaces. Each subspace is associated with an atom and electron occupancy divided between the subspaces. In the NBO procedure this subspace is divided equally into bonding and anti-bonding orbitals. In the Kitaura Morokuma procedure, not all the atomic orbitals are equally divided into bonding and anti-bonding orbitals. In the case of the water dimer, 8 of the 12 valence orbitals are divided in this way but the other 4 orbitals are treated with a Schmidt orthogonalization procedure that results in antibonds with unusually low occupancy. The net result for this is that electron density gets sequestered into bonding orbitals instead of antibonds and charge-transfer contributions misclassified as electrostatic contributions.[45]

One other method to calculate charge transfer effects is the block-localized wavefunction (BLW) method.[46,47] The main difference between this method and NBO is that the localized wavefunction constructed by the NBO

method is not optimized to self-consistency while the block-localized wavefunctions are. Again compared to the BLW method, NBO has been shown to overestimate charge-transfer effects. However, it has been shown that NBO produces similar qualitative results with the BLW method. Thus while caution must be exercised in treating standalone NBO charge-transfer numbers, NBO can be a robust method for deriving qualitative trends. [47] However, the most compelling reason to use NBO remains the fact that NBO can quantify delocalization interactions between individual orbitals. Both the Morokuma and BLW methods only quantify total charge transfer within molecular species.

Natural Resonance Theory

One important offshoot of NBO is the natural resonance theory. Here the density matrix is represented as a sum of individually weighted density matrices, equation 18.

$$D_{true} = \sum_{\alpha} w_{\alpha} D_{\alpha} \text{ [eqn 18]}$$

where w_{α} represents the % weight of the resonance form described by density matrix D_{α}. This method allows a quantification of which resonance form dominates a population and can be useful in understanding the underlying chemical origin of species dominating reaction pathways.

"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."

Introduction

       Phosphagens (N-phosphorylguanidines) are a group of small, phosphorylated biomolecules that participate in phosphoryl transfer reactions [48-51], as shown in Figure 3.1.
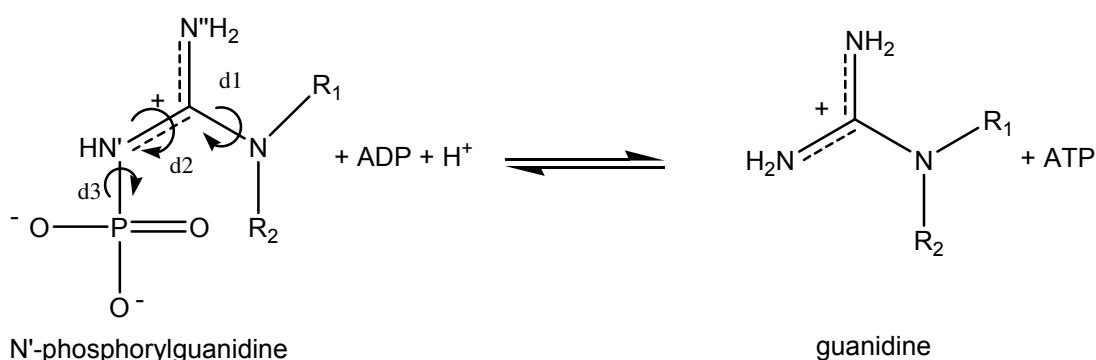
Figure 3.1. A schematic of the phosphagen kinase reaction as modeled with N'-phosphorylguanidine. The designation of N, N' and N'' for the guanidinium nitrogens is used throughout the text. The R_1 group of N'-phosphorylguanidine, which does not participate in the reaction, is modeled as a methyl group. In most N'-phosphorylguanidines, or phosphagens, R_2=H except for creatine where R_2=CH_3[1].

       Phosphoryl transfer reactions that break and form high-energy phosphoester bonds have been recently referred to as "the centerpiece of biochemical processes" due to their fundamental roles in metabolism and signaling [1,52-55]. The reversible transfer of the phosphoryl group from N-phosphorylguanidines to ADP is crucial in maintaining steady ATP concentrations. ATP is the currency from which most biological processes derive all their energy needs and this reversible phosphoryl transfer is essential in

tissues subject to rapid fluctuations in energy demands, such as the heart muscle [51] and nervous tissue [56].

While crystallography [17,57-61], biochemical kinetics [59,61-64] and thermodynamic studies [65,66] have all furthered the understanding of phosphagen kinase enzymes that catalyze the reaction in biological systems, critical molecular level aspects surrounding the chemistry of N-phosphorylguanidine substrates remain unknown. Among the needed studies is a clear understanding of the molecular origin of the labile or high-energy N'—P bond between the phosphoryl and guanidine groups [67]. The weakness of the N'—P bond is essential to the physiological role of phosphagens, where the exothermic cleavage of the N'—P bond is coupled to the endothermic formation of the O—P bond, thus enabling the favorable formation of ATP. The determinants of N'—P bond lability control the capacity of phosphagens in fulfilling their role in providing phosphoryl groups in ATP formation, and to date have not been clearly defined nor quantified.

Both N'—P and O—P bonds have been described as labile or high-energy bonds owing to the fact that the hydrolysis of these bonds is accompanied by relatively large negative standard free energy change, $\Delta G°$. It is important to note that the term "high energy bond" is not the same as "bond energy", which is defined as the energy required to break, not hydrolyze, a covalent bond [6]. The standard free energy change of ATP hydrolysis is ~ -7.3 kcal/mol and that of N'-phosphorylcreatine is ~ -10.3 kcal/mol, indicating a preference of O—P over N'—P bond formation. However under more complex physiological conditions, there is little difference in the overall free energy change [6,68]. Because both compounds exhibit relatively large negative free energy changes during hydrolysis [6], the chemical rationalization of the high-energy character of the N'—P bond in N-phosphorylguanidines has been assumed implicitly to be similar to that of the O—P bond in ATP[6,68]. Because phosphoryl transfer involves the breaking or formation of high-energy phosphoryl bonds, the chemistry of these bonds has a profound effect on mechanism and catalysis[1].

The high-energy character or weakness of the O—P phosphoanhydride bonds in ATP has been traditionally attributed to three main factors. Firstly, resonance stabilization in the phosphoester reactant is less favorable than in the hydrolyzed products causing the O—P bond to be weakened [6,12,68,69]. Protonation states directly impact resonance stabilization and control of mechanism in phosphate monoesters [70,71,72]. Weaker resonance stabilization in the phosphodiester has been rationalized as an "opposing resonance effect" brought about by competition between adjacent, resonance stabilized, phosphoryl groups for the same lone pair on the bridging oxygen atom [6,10,12,69,73,74]. Secondly, electrostatic repulsion between negatively charged oxygen atoms on the adjacent phosphoryl groups destabilizes the esterified form and should lengthen the O—P bond [6,68]. The phosphagen guanidinium and phosphoryl groups are oppositely charged so electrostatic repulsion does not contribute to the lability of the N'—P bond. Finally, solvation of ATP has been shown to be less favorable than that of its hydrolysis products [6,14,68]. Apart from solvation, these arguments have been qualitative.

Resonance factors, including opposing resonance theory, have been used to explain N'—P bond lability in N-phosphorylguanidines due to energetic similarity with O—P phosphoanhydride bonds in ATP [6,12]. However, there are no reports in the literature that directly examine the resonance factors governing the underlying weakness of the N'—P bond and subsequently its ability to affect phosphoryl transfer reactions or mechanism. One reason for this is that the weakness of the N'—P bond itself makes phosphagens in free acid forms difficult for experimentation. Indeed to the best of our knowledge, geometric parameters of phosphagens in isolation have not been reported.

Crystallographic studies of a phosphagen kinase enzyme, transition state analog complex [17] suggests that the enzyme stabilizes a transient, hitherto unanticipated quaternary form of the phosphorylated N' nitrogen. In phosphagen kinase enzymes, electrostatic calculations suggest that it is entirely conceivable that alternate protonated forms may exist [59,60]. Because protonation states influence conjugation and resonance, the reported electrostatic calculations

emphasize the need for detailed studies of phosphagen structure and their determinants of stability. However, even if only one form predominates in biology, the aim of using the different forms in this is to probe any possible association between N'—P bond length and resonance through systematic investigation.

Due to the physiological importance and lack of understanding of the high-energy N'—P bond in significant biochemical reactions, a systematic study of N'—P bond lability has been undertaken. Computational methods have been utilized to study resonance and stereoelectronic effects of different protonation states upon the chemistry of the N'—P bond. Electronic structure calculations with a variety of basis sets and quantum chemical methods have been employed to investigate the structure and stability of the N'—P bond. Electronic structure methods with natural bond order analysis provide the necessary atomistic and stereoelectronic detail to identify and quantify the determinants of N'—P lability.

<u>Computational Methods</u>

All electronic structure calculations were carried out with the Gaussian program [75,76] using the computational resources at the FSU School for Computational Science and Information Technology (CSIT), the Center for Computational Sciences [77] at Duquesne University and the Pittsburgh Supercomputer Center [78].

To identify reasonable, low-energy gas-phase conformations, the protonated systems were subjected to grid searches along the three main dihedrals, C-N-C-N'' (d1), N-C-N'-P (d2) and C-N'-P-O (d3), as shown in Figure 3.1. The dihedral angles were scanned in 30º increments from 0 to 180º using the PM3[79,80] semiempirical method. PM3 was chosen because of its ability to provide stationary points as a starting point for further optimizations at higher levels of theory for a variety of phosphorus and nitrogen containing compounds [81,82]. The energy minimized structures were located using Hartree-Fock (HF), density functional theory (DFT) and second-order many-body Møller-Plesset [83] theory (MP2). Specifically, DFT was implemented by using Becke's three-parameter hybrid (exchange) functional [84] with gradient corrections provided by the Lee, Yang, and Parr [85] (B3LYP). Acknowledging the lack of dispersive forces in DFT

methods [86], MP2 optimizations were carried out on all structures to serve as a point of verification and test of DFT. To further validate the MP2 and B3LYP geometries and energies, two selected structures were optimized using coupled-cluster theory (CCSD) [87-89] and quadratic CI (QCISD) [90,91]. The Pople style 3-21G [92-94], 6-31G(d) [95-97], 6-31+G(d) [95-98] and 6-311++G(d,p) [98-100] basis sets were utilized. Contributions due to thermal, vibrational, rotational and translational motions, including zero-point energies, were included separately by standard statistical mechanical procedures available in Gaussian. Frequency analysis has been used to confirm all stationary points as minima or transition structures and provide thermodynamic and zero-point energy corrections [101].

<u>Results and Discussion</u>

A two-step approach is used to identify the stereoelectronic effects responsible for the lability of the N'—P bond of N-phosphorylguanidines. First, through electronic structure calculations, clearly defined trends are established between the different protonation states and N'—P bond lengths, with N'—P bond lengths being the physical characteristic associated with N'—P lability. Second, using an NBO analysis of the different protonated forms, N'—P bond lengths are shown to depend upon the delocalization and stereoelectronic interactions of phosphoryl and guanidinium groups.

Phosphagen geometries and energies. N-methyl-N'-phosphorylguanidine is used as a model compound representative of phosphagens. The model contains atoms common to all phosphagens, but in this model, $R_1 = CH_3$ where the substituent in natural phosphagens is one of several groups from acetate to ethyl phosphoserine [102-104]. A total of five protonation sites (Figure 3.2) of N-methyl-N'-phosphorylguanidine, three of which are chemically identical (sites 1-3), have been investigated.

Of the nine possible geometries investigated (Figure 3.3), 16 low-energy configurations, two conformations each for 1b-3c, have been selected from the grid searches at the PM3 level for further investigation. The difference between the first and second lowest energy structure is in all cases an 180° rotation about the dihedral labeled d2 in Figure 3.1. The energy difference between the two low

energy structures is less than 1.0 kcal/mol at the PM3 level. Where a third low conformation is found, the energy difference between the second and third lowest energy structures is at least 6.0 kcal/mol, so these higher energy conformations are not further considered.
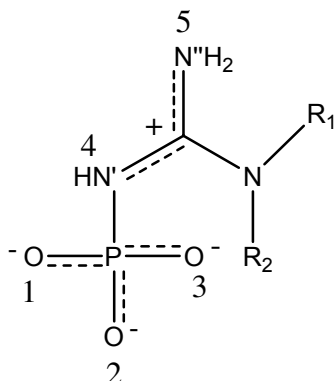
Figure 3.2. N-methyl-N'-phosphorylguanidine with protonation sites labeled 1 through 5. The designation of N, N' and N'' used in the text is shown.

The remaining 16 low-energy structures have been further energy minimized at levels of theory ranging from HF/3-21G to B3LYP/6-311++G(d,p). At the B3LYP/6-311++G(d,p) level, the energy difference between the d2 rotamers increases to at least 7.0 kcal/mol. Therefore, only the lower energy E-forms have been further optimized using MP2/6-311++G(d,p). The computed N'—P bond lengths are given in Tables 3.1 and 3.2. Relative energies of the final eight structures are given in Table 3.3. There are no imaginary frequencies for any of the eight structures indicating that each structure is a true minimum.

Validation of acceptable quantum chemical methods and basis sets is complicated by the lack of prior theoretical calculations and the limited experimental information available. The only experimental structure comes from a crystal of the hydrated sodium salt of N-phosphorylcreatine [105]. Its relevance here is limited by the substitution of $R_2=CH_3$ in creatine instead of $R_2=H$ in other phosphagens. This affects the C—N bond length and delocalizing interactions involving C—N' and C—N''. Secondly, the coordination of the Na^+ counterions to the phosphoryl group affects the O—P bond lengths relative to the free acid forms from the computations.
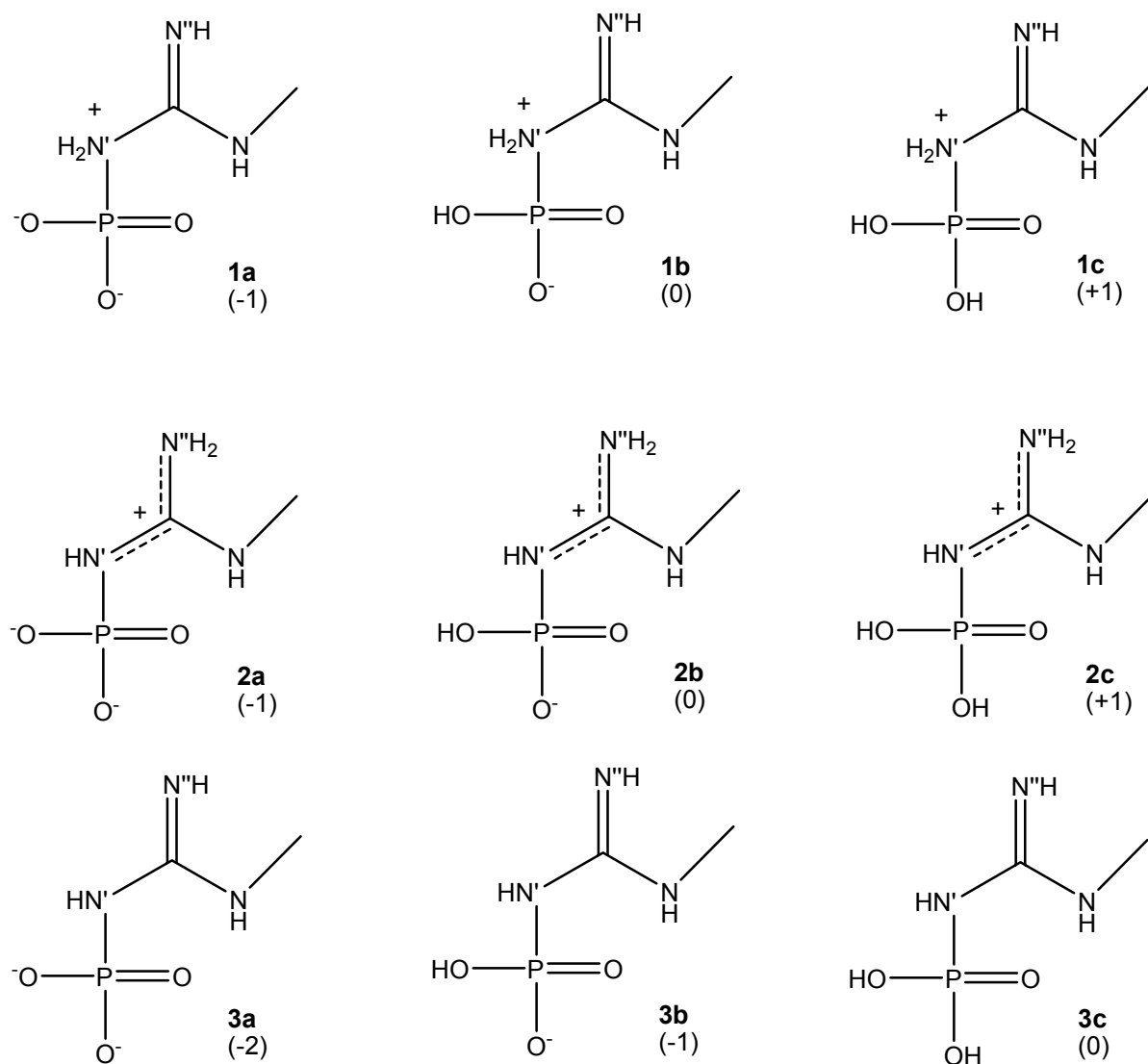
Figure 3.3. The nine structures studied in this work. Overall charges are given in parentheses.

The computations are validated by comparing structures and energies with results from higher levels of theory. The CCSD/6-311++G(d,p) and QCISD/6-311++G(d,p) levels of theory have been selected as the standards, since they have been shown to produce accurate structures and relative energies [106-108]. The N'—P bond of both isomers of 1a dissociates during energy minimization attempts using all levels of theory employed. Isomers of 2a also dissociate during energy minimization using MP2/6-311++G(d,p), CCSD/6-311++G(d,p) and QCISD/6-311++G(d,p). However, all other levels of theory give a stationary point for 2a. Even though minima are located, it is clear that the

lower levels of theory do not predict a stabilized structure, since the N'—P bonds are longer than 2.0 Å (Table 3.1).

Table 3.1: N'—P bond lengths calculated at different levels of theory.

	HF/ 6-31G(d)	HF/ 6-31+G(d)	HF/ 6-311++G(d,p)	B3LYP/ 6-31G(d)	B3LYP/ 6-31+G(d)	B3LYP/ 6-311++G(d,p)	MP2/ 6-311++G(d,p)
1b	1.947	1.940	1.938	2.007	1.998	2.007	1.979
1c	1.802	1.803	1.796	1.836	1.837	1.830	1.823
2a	2.055	2.004	2.004	2.014	2.021	2.010	
	nd						
2b	1.796	1.788	1.784	1.831	1.821	1.821	1.816
2c	1.701	1.699	1.694	1.722	1.721	1.717	1.710
3a	1.845	1.831	1.825	1.883	1.870	1.870	1.856
3b	1.718	1.715	1.711	1.742	1.740	1.737	1.737
3c	1.658	1.658	1.654	1.676	1.677	1.673	1.673

With differences in 2a between levels of theory, it was important to validate the MP2/6-311++G(d,p) and B3LYP/6-311++G(d,p) methods for the other structures. The energy-minimized structures of 2b are compared with those computed at CCSD/6-311++G(d,p) and QCISD/6-311++G(d,p). As seen in Table 3.2, the MP2 and B3LYP N'—P bond lengths show little change with the higher levels of theory. Using the 6-311++G(d,p) basis set, the difference in the N'—P bond lengths between MP2 and CCSD is 0.008 Å where the difference is 0.006 Å between MP2 and QCISD. Overall across selected bond lengths, the difference when comparing MP2 with CCSD and QCISD is 0.004 ± 0.003 Å and 0.004 ± 0.002 Å, respectively. Consequently, MP2/6-311++G(d,p) is found to produce structures in solid agreement with both CCSD/6-311++G(d,p) and QCISD/6-311++G(d,p). B3LYP/6-311++G(d,p) is also in strong agreement with the higher levels of theory (Table 3.2).

Table 3.2: Structural comparison of 2b at high levels of theory.

Structure 2b	B3LYP/ 6-311++G(d,p)	MP2/ 6-311++G(d,p)	CCSD/ 6-311++G(d,p)	QCISD/ 6-311++G(d,p)
N—P	1.821	1.816	1.808	1.810
P—O1	1.636	1.629	1.624	1.624
P—O2	1.477	1.476	1.469	1.471
P—O3	1.508	1.507	1.498	1.500
C—N'	1.335	1.335	1.333	1.334
C—N''	1.362	1.369	1.369	1.371
C—N	1.329	1.327	1.327	1.328

Average error between B3LYP and CCSD = 0.007 ± 0.003 Å; Average error between B3LYP and QCISD = 0.007 ± 0.004 Å;
Average error between MP2 and CCSD = 0.004 ± 0.003 Å; Average error between MP2 and QCISD = 0.004 ± 0.002 Å.

Fully optimized structures together with N'—P bond lengths at both the B3LYP/6-311++G(d,p) and MP2/6-311++G(d,p) are shown in Figure 3.4. B3LYP produces geometric results that are close to MP2 when using the 6-311++G(d,p) basis set for all structures (Table 1 and S1) except 2a (where B3LYP predicts a stationary point). In fact, the computed average energy difference is 0.5 ± 0.2 kcal/mol between MP2/6-311++G(d,p)//MP2/6-311++G(d,p) and MP2/6-311++G(d,p)//B3LYP/6-311++G(d,p) across all computed structures (Table 3.3). Most importantly, all levels of theory show bond lengths that are correlated, i.e. all bond lengths change from structure to structure in similar ways. For example, the N'—P bond length always decreases in the following order: 1b > 3a > 1c > 2b > 3b > 2c > 3c for all levels of theory.

Stereoelectronic and anomeric effects. The weakness of the N'—P bond has been traditionally explained by the increased resonance stability of the dissociated phosphoryl and guanidinium groups compared to the N-phosphorylguanidine product [6,12,68,69]. The framework for increased resonance stability has been given by opposing resonance theory[6,12,69], where the phosphoryl and guanidinium groups compete for the same lone pair on the bridging nitrogen, N', giving a weakened N'—P bond (Figure 3.5).

Table 3.3: Relative energy differences (between structures) calculated by the various methods including scaled Z.P.E. The scale factor used for the ZPE was 0.9135 for HF/6-31G(d), 0.9163 for HF/6-31+G(d) and HF/6-311++G(d,p), 0.9806 for all B3LYP calculations and 0.9748 for MP2 calculations [80].

Level of theory	1b-2b	1c-2c	1b-3c	2b-3c	2a-3b
HF/6-31G(d)	25.6	55.1	19.8	-5.8	28.5
HF/6-31+G(d)	25.4	54.2	19.1	-6.2	27.5
HF/6-311++G(d,p)	26.2	54.7	23.0	-3.3	29.9
B3LYP/6-31G(d)	20.7	46.4	11.6	-9.1	20.1
B3LYP/6-31+G(d)	20.3	45.6	10.6	-9.6	19.8
B3LYP/6-311++G(d,p)	21.2	46.0	13.6	-7.6	20.7
MP2/6-311++G(d,p)	17.3	42.4	12.6	-4.7	nd
MP2/6-311++G(d,p)//HF/6-311++G(d,p)	18.1	43.6	13.8	-4.3	25.0
MP2/6-311++G(d,p)//B3LYP/6-311++G(d,p)	18.1	42.7	12.8	-5.3	23.4

Average error between MP2 and B3LYP = 2.9 ± 0.9 kcal/mol; Average error between HF and B3LYP = 8.3 ± 3.4 kcal/mol;
Average error between MP2//MP2 and MP2//B3LYP = 0.5 ± 0.2 kcal/mol.
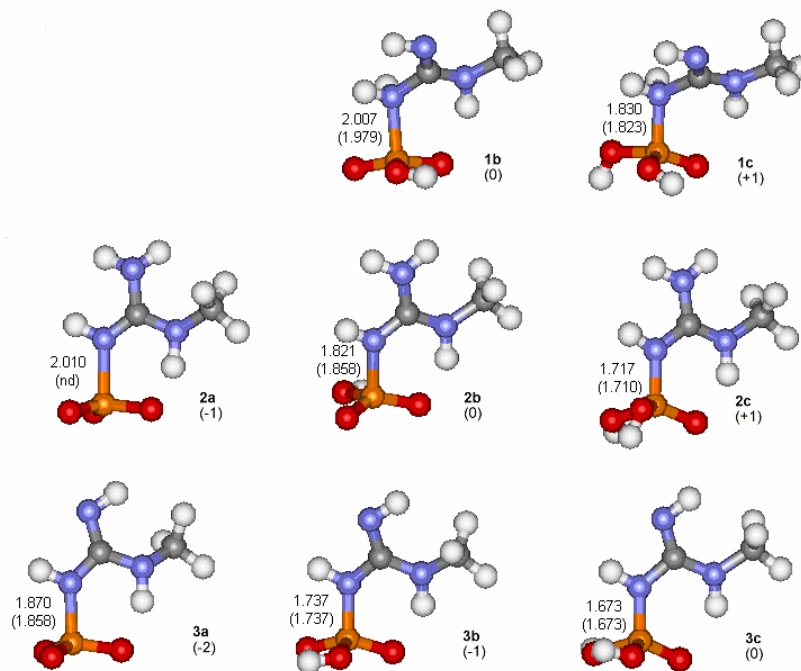
Figure 3.4. Eight optimized structures of N-methyl-N'-phosphorylguanidine. N'—P bond lengths in Å are given at B3LYP/6-311++G(d,p) and MP2/6-311++G(d,p) (in parentheses).
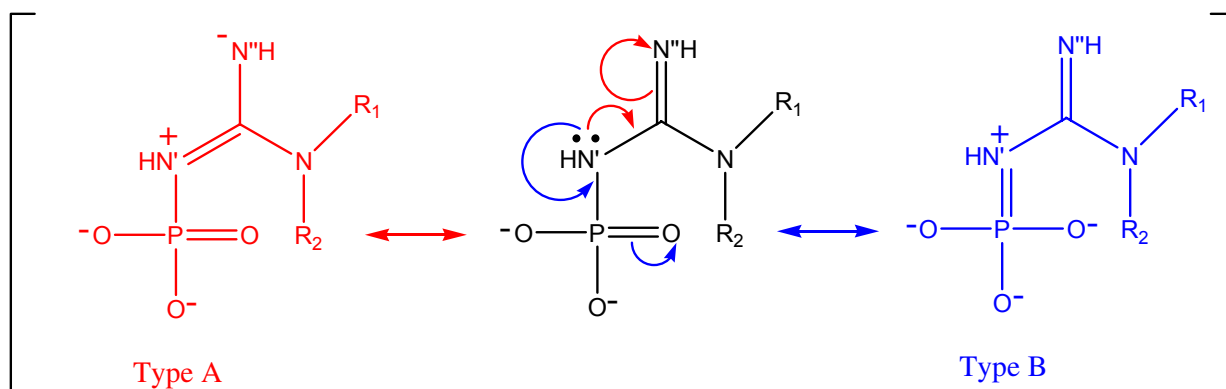
Figure 3.5. In opposing resonance theory, there are two types of resonance forms, Type A and Type B. It has been proposed that both types compete for the same lone pair on the bridging nitrogen, N'. As a result, the N'—P bond is weakened[18-21]. In NBO nomenclature, the delocalization interaction responsible for Type A is identified as n(N') → σ*(C—N'') or n(N') → σ*(C—N) and Type B is identified as a n(N') → σ*(P—O) orbital interaction.

By changing the protonation states of N-methyl-N'-phosphorylguanidine, investigation of the relationship between increased resonance at the phosphoryl and guanidinium groups and N'—P bond lengths, as well as testing opposing resonance theory, is possible. By examining the amount of electron delocalization from the lone pair on the N' to the guanidinium and phosphoryl groups using an NBO analysis, we obtain a quantitative estimate of the extent to which competition for the lone pair on N' is correlated with N'—P bond lengths. Specifically, through second-order perturbation methods, an NBO analysis quantifies the interactions between Lewis and non-Lewis type NBOs, showing how much the actual delocalized system deviates from the idealized Lewis structure. Second-order perturbation energy terms provide a quantitative estimate of the delocalizing interactions. Therefore, NBO analysis provides a convenient method to identify and relate specific delocalization interactions with N'—P bond lengths.

Since the identity of the delocalizing interactions responsible for the trend in N'—P bond lengths are of interest, we focus on interactions with σ(N'—P) as a donor or with σ*(N'—P) as an acceptor. Note that electron density loss from σ(N'—P) and electron density gain in σ*(N'—P) both lead to longer, weaker N'—P bonds. The delocalization interaction responsible for the resonance form Type A is identified as n(N') → σ*(C—N'') or n(N') → σ*(C—N), while that of Type B is

given by n(N') → σ*(P—O), as illustrated in Figure 3.5. We also focus on these interactions to investigate the correlation between N'—P bond length and the magnitude of such interactions, in order to test opposing resonance theory. These findings are presented below.

Role of oxygen lone pairs in destabilizing the N'—P bond. The six structures, 2a through 3c, show distinct variation in N'—P bond lengths (Table 3.1, Figure 3.4). Two trends can be seen in the calculations where increased phosphoryl protonation results in shorter, stronger N'—P bonds and protonation at the unsubstituted N'' nitrogen leads to longer, weaker N'—P bonds.

Comparing structures isoprotonic at the guanidinium group, we find that increasing protonation at the phosphoryl group leads to shorter N'—P bonds, where 2a > 2b > 2c and 3a > 3b > 3c (Table 1, Figure 3.4). Even for tetrahedral N', structures 1b and 1c showed similar N'—P bond length trends where 1b > 1c. This is true at all levels of theory tested (Table 1). From a qualitative perspective, these results clearly show that decreased phosphoryl resonance leads to shorter N'—P bonds, as expected [6,68,70].

The increase in N'—P bond lengths is strongly correlated to the total magnitude of N'—P bond weakening n(O) → σ*(N'—P) interactions. The correlation holds regardless of method used (Table 3.4 and Figures 3.6 and 3.7) or the presence of solvation (Figure 3.7). R^2 values exceed 0.9 in all cases.

Electron density delocalization can be considered as interference between waves. As such destructive interference occasionally results. Such destructive interference called anti-cooperative interactions, are destabilizing. Since in this case, there is more than one lone pair donating into there is a possibility that destructive interference between the lone pair delocalizations occur. This would lead to different values in overall anomeric energy that may not give the correlations given in Figure 3.6 and 3.7. To check that the correlations were correct and that the interactions were not anti-cooperative, an NBO deletion analysis comparing the total E(2) energy values to the simultaneously deleted interactions [23,109-114] was performed. With all structures, there is a slight positive cooperativity (Table 3.4). Nevertheless, the correlation between the total n(O) →

σ*(N'—P) energy terms and N'—P bond lengths made does not change whether E(2) or deletion energies are used (Figure 3.6). The correlation of the simultaneously deleted energies with the N'—P bond lengths is insignificantly better with R^2 = 0.989 compared to R^2 = 0.988 for the sum of the E(2) values. Due to the insensitivity of the computed values to changes in solvent and method, E(2) values at the B3LYP/6-311++G(d,p)// B3LYP/6-311++G(d,p) level in vacuum is discussed throughout the manuscript. To probe the effect of solvent on the structures, optimizations were performed on all structures using the PCM method in toluene (dielectric constant, 2.379) and water (dielectric constant, 78.39). The structures show that with increasing dielectric constant, the N'—P bond length shortens with no difference in the trends between N'—P bond lengths and anomeric energies (shown in Figure 3.7).

Table 3.4: Deletion analysis of the anomeric effect. E_{total} is the sum of individual deletions, and E_{sum} is the interactions deleted simultaneously at the HF/6-311++G(d,p)//B3LYP/6-3-11++G(d,p) and HF/6-311++G(d,p)//MP2/6-3-11++G(d,p)optimized structures. All values are in kcal/mol.

	2a	2b	2c	3a	3b	3c
HF/6-311++G(d,p)//B3LYP/6-311++G(d,p)						
E_{sum}	78.7	57.1	40.1	60.8	46.8	39.2
E_{total}	66.3	50.0	36.7	54.3	42.0	35.6
HF/6-311++G(d,p)//MP2/6-311++G(d,p)						
E_{sum}	nd	57.6	40.3	59.3	46.8	38.8
E_{total}	nd	50.3	36.8	53.2	42.0	35.8
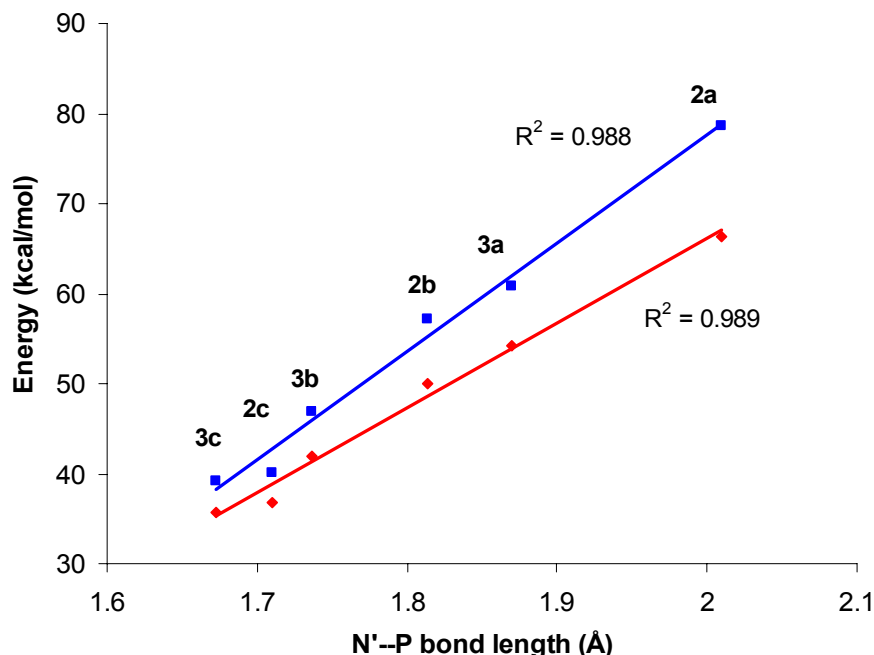
Figure 3.6. The graph shows the correlation between the energy values of the interactions of the n(O) → σ*(N'—P) type with N'—P bond lengths using energies (red line) deleted simultaneously and sum of the E(2) energies (blue line). The level of theory used here is HF/6-311++G(d,p)//B3LYP/6-311++G(d,p).

The sum of the interactions is presented. This is comprised of three reinforcing n(O) → σ*(N'—P) interactions from each of the three phosphoryl oxygen atoms (six to eight lone pairs). The individual contributions from each interaction is approximately 20–30 kcal/mol when fully deprotonated (1a-3a). Our computed values are consistent with other n̩(O) values that have been previously published [24,25].

The n(O) → σ*(N'—P) interaction (Figure 3.8) is a generalized anomeric interaction of the Lp-X-A-Y variety where "Lp" represents a lone pair, "X", any heteroatom, "A" an electropositive atom and "Y" and electronegative atom [22-25]. The generalized anomeric effect has been well documented as being responsible for a wide variety of effects most notably in molecular conformation [115-117].
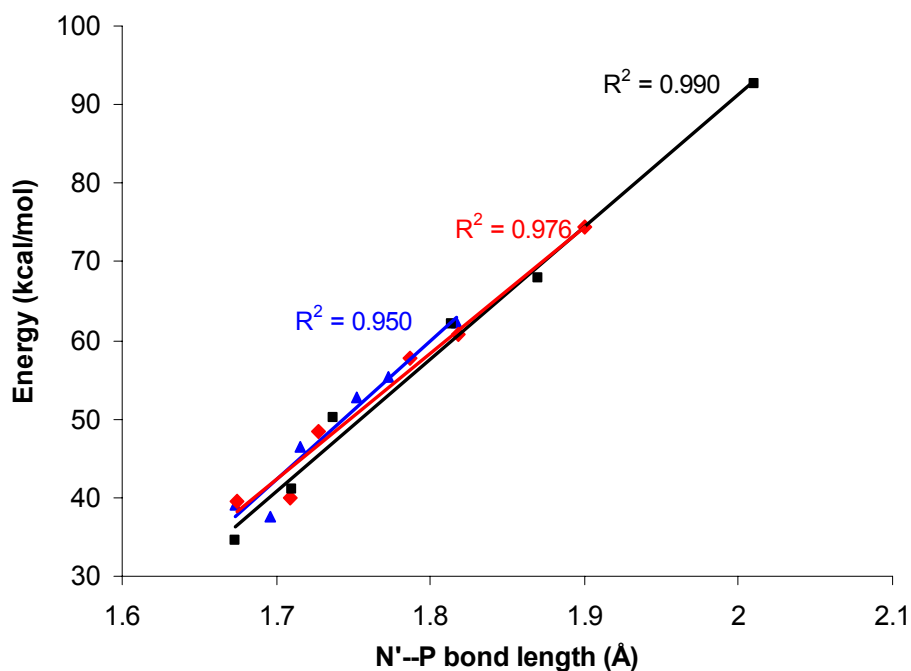
Figure 3.7. The graph shows the correlation between the energy values of the interactions of the n(O) → σ*(N'—P) type with N'—P bond at the B3LYP/6-311++G(d,p)// B3LYP/6-311++G(d,p) level. The black line shows the gas phase correlation, the red line at a dielectric of 2.379 and the blue line 78.39. Each set (gas phase, dielectric=2.379 and dielectric=78.39) include, in ascending order of bond lengths, structures 3c, 2c, 3b, 2b, 3a and 2a.
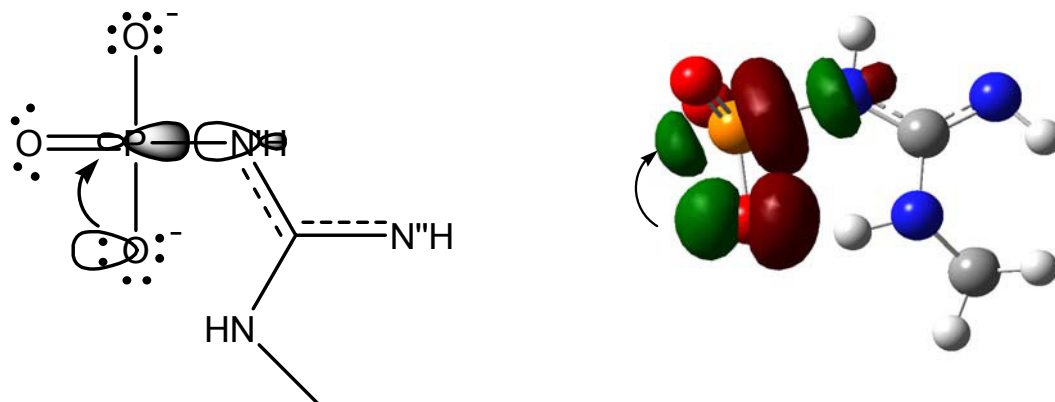
Figure 3.8. A schematic and orbital diagram of the anomeric interaction in 3a between n(O) and σ*(N'—P).

The computations provide an explanation on why increased phosphoryl protonation results in shorter N'—P bonds. A phosphoryl group has three lone pairs on each of the two negatively charged oxygen atoms and two lone pairs on the neutral oxygen atom (Figure 3.8). Successive protonation removes lone pairs of electrons from a possible anomeric interaction. The anomeric effect is further reduced by increasing the energy difference between the average n(O) lone pair

energy and σ*(N'—P), as shown in Table 3.5. As protonation at the phosphoryl group increases (2a < 2b < 2c and 3a < 3b < 3c), so does the n(O) → σ*(N'—P) energy difference (2a < 2b < 2c and 3a < 3b < 3c).

Table 3.5: Energy values of the oxygen lone pairs and σ*(N'-P). As protonation on the phosphoryl group increases, the average energy of the oxygen lone pairs decrease as do the σ*(N'-P) energies. However, the energy gap between the oxygen lone pairs and σ*(N'-P) orbital increases leading to a more pronounced anomeric n(O) → σ*(N'—P) interaction. All values are in a.u.

Level of Theory	Average Structure	n(O) energy	Energy σ*(N'-P) energy	difference
B3LYP/6-311++G(d,p)	2a	-0.25011	0.58185	0.83196
	2b	-0.61305	0.37306	0.98611
	2c	-0.82806	0.25956	1.08762
	3a	-0.26191	0.69213	0.95404
	3b	-0.47168	0.60857	1.08025
	3c	-0.68679	0.46401	1.15080
MP2/6-311++G(d,p)	2b	-0.61572	0.37406	0.98978
	2c	-0.83144	0.26587	1.09731
	3a	-0.26188	0.69487	0.95675
	3b	-0.47299	0.60497	1.07796
	3c	-0.68971	0.45981	1.14952

As discussed, protonation of the phosphoryl group modulates the number and quality of the n(O) → σ*(N'—P) anomeric interactions, which is the primary factor in determining the N'—P bond length. Structure 2a has the same phosphoryl protonation state as 3a, 2b as 3b and 2c as 3c. Therefore, it is expected that the magnitude of n(O) → σ*(N'—P) should be equal for the pairs of structures 3a and 2a, 3b and 2b, and similarly 3c and 2c. The secondary factor found involves the guanidinium protonation state. Consider structures from series 2, which are protonated at N'' to give a delocalized guanidinium. As such, N' becomes more sp^2-like for structures 2a-2c compared to 3a-3c. The change in hybridization of N' stabilizes the σ*(N'—P) orbital allowing for more efficient overlap with n(O) lone pairs yielding a stronger anomeric effect for series 3. This is seen in Table 3.5. where the difference in energies between the average lone pair energy and σ*(N'—P) is less in series 2 compared to 3 (2a < 3a, 2b < 3b and 2c < 3c). Therefore, the protonation state changes for both series 2 and 3 are related and have the same influence on the anomeric effect giving a linear response.

The ranking of the N'—P bond lengths is as expected with the primary (phosphoryl protonation) and secondary (guanidinium protonation) factors in operation, 3c < 2c < 3b < 2b < 3a < 2a. Structure 3c has the shortest N'—P bond, since the phosphoryl group is fully protonated and the guanidinium is not protonated, minimizing the anomeric effect. Structure 2c has the second shortest N'—P bond, since with the same phosphoryl protonation state, the N'' guanidinium is protonated increasing the anomeric interaction. The pattern continues through all six structures.

The results explain why a stationary point could not be located for structure 2a at the higher levels of theory. We find a large n(O) → σ*(N'—P) energy interaction for structure 2a. It is possible that the N'—P bond fails to form in 2a at the MP2/6-311++G(d,p) level, because of the bond weakening anomeric interaction. The lower levels of theory may not have properly described this phenomenon and produced a stationary point.

Role of N—P bond as a donor; interactions with guanidinium. Comparing structures that are isoprotonic at the phosphoryl group, but either singly or doubly protonated at N'', we find in terms of N'—P bond lengths that 2a > 3a, 2b > 3b and 2c > 3c (Table 3.1, Figure 3,4). Doubly protonated N'' forms display longer N'—P bond lengths. Again this is true at all levels of theory. Any nitrogen atom on a guanidinium group with only one proton effectively locks the lone pair on that nitrogen into a double C=N bond, preventing it from participating in delocalization interactions involving the guanidinium group. Two protons on N'' makes it more sp^3-like, allowing more conformational freedom and possible delocalization with the guanidinium system.

The NBO results also show σ(N'—P) → σ*(C—N) and σ(N'—P) → σ*(C—N'') interactions that could contribute to the weakening of the N'—P bond (Figures 3.9 and 3.10). The total magnitude of the N'—P bond weakening interactions are correlated with N'—P bond lengths. The E(2) and Edel magnitudes are 2a > 3a, 2b > 3b and 2c > 3c mirroring differences in N'—P bond lengths (Tables 3.4, Figure 3.9). This can be rationalized by the fact that deprotonation of N'' locks the lone pair on N'' into a double bond at C=N'' making it difficult for other delocalized

forms to exist [54] in series 3 structures relative to series 2. As the N'—P bond is a source of electrons for guanidinium delocalization, it is lengthened as it is weakened with increasing σ(N'—P) → σ*(C—N) and σ(N'—P) → σ*(C—N'') interactions in series 2. The correlation (R^2 = 0.994) between the sum of σ(N'—P) → σ*(C—N) and σ(N'—P) → σ*(C—N'') interactions and N'—P bond lengths is clear, but these stabilizing interactions are weaker than the generalized anomeric interaction n(O) → σ*(N'—P) previously discussed.
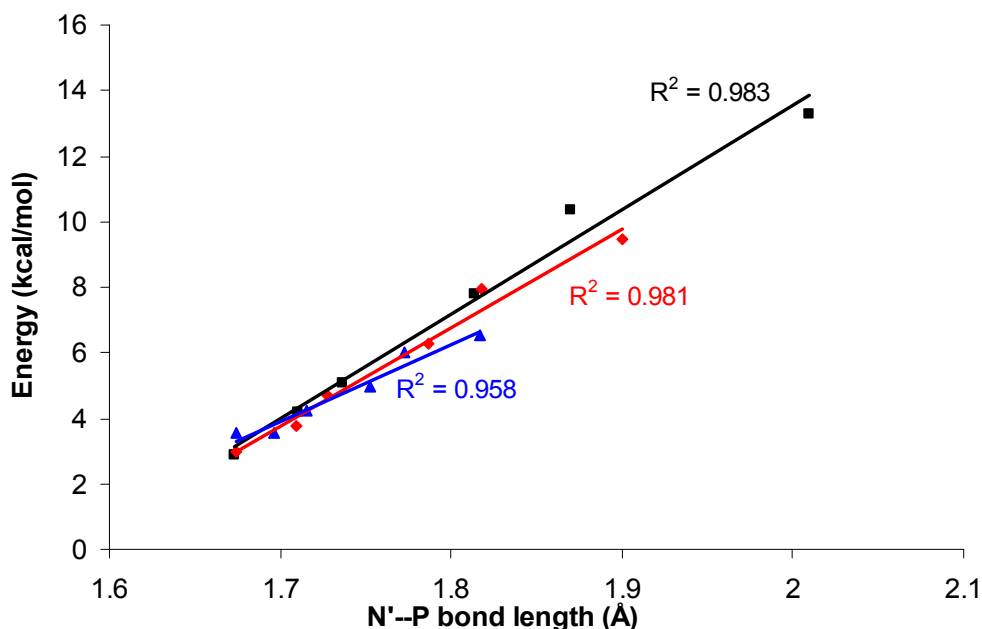
Figure 3.9. The correlation between N'—P bond lengths and the sum of the σ(N'—P) → σ*(C—N) and σ(N'—P) → σ*(C—N'') interaction energies at the B3LYP/6-311++G(d,p)//B3LYP/6-311++G(d,p) level. The black line shows the gas phase correlation, the red line shows the correlation at a dielectric of 2.379 and the blue line a dielectric of 78.39. Each set (gas phase, dielectric=2.379 and dielectric=78.39) include, in ascending order of bond lengths, structures 3c, 2c, 3b, 2b, 3a and 2a.

Being of the same protonation state at the guanidinium, 2a, 2b and 2c as well as 3a, 3b and 3c are expected to have almost equal σ(N'—P) → σ*(C—N) and σ(N'—P) → σ*(C—N'') interactions. The fact that the σ(N'—P) → σ*(C—N) and σ(N'—P) → σ*(C—N'') interactions decrease in order of 2a > 2b > 2c and 3a > 3b > 3c according to N'—P bond lengths may be explained by a better overlap between the σ(N'—P) and σ*(C—N) and σ*(C—N'') orbitals and stronger stereoelectronic effect in forms with less protonation at the phosphoryl group, as shown in Table 3.6. The trend in the energy difference between the σ(N'—P) and

σ*(C—N) and σ*(C—N'') is 2a < 2b < 2c and 3a < 3b < 3c, thus explaining the linearity seen for all structures in Figure 9.

Table 3.6 : Energy values of σ (N'-P) and σ*(C-N) as well as σ*(C-N''). All values are in a.u.

Level of theory	Structure	Energy σ (N'—P)	Energy σ*(C-N) σ*(C-N'')	Energy difference between σ N'--P and σ*(C-N'')	(σ*C-N)
B3LYP/6-311++G(d,p)	**2a**	-0.31439	0.61421	0.49855	0.92860	0.81294
	3a	-0.45140	0.99038	0.95083	1.44178	1.40223
	2b	-0.85665	0.75233	0.73937	1.60898	1.59602
	3b	-0.74212	0.87692	0.89493	1.61904	1.63705
	2c	-1.12617	0.56534	0.61842	1.69151	1.74459
	3c	-0.98881	0.74757	0.86853	1.73638	1.85734
MP2/6-311++G(d,p)	**2b**	-0.72771	0.75227	0.66135	1.47998	1.38906
	3b	-0.86125	0.89237	0.86053	1.75362	1.72178
	2c	-0.97959	0.61924	0.55944	1.59883	1.53903
	3c	-1.13046	0.73638	0.79137	1.86684	1.92183

a

b
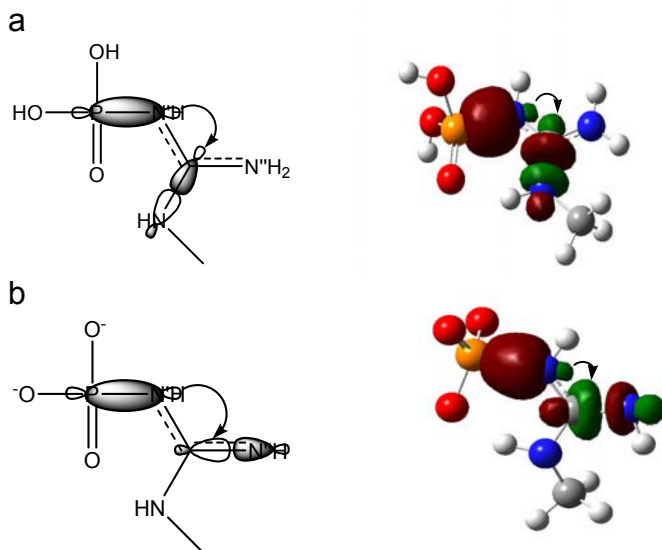
Figure 3.10. Stereoelectronic interactions a) σ(N'—P) → σ*(C—N) in 2c and b) σ(N'—P) → σ*(C—N'') in 3a.

N-methyl-N'-phosphorylguanidine with a quaternary nitrogen at N'. The calculations indicate that when N' is quaternary, the N'—P bond is weakened. Structures 1b and 1c both show a quaternary nitrogen at the N' position. The N'—P bonds of 1b and 1c are longer than their isoprotonic counterparts 2b and 2c that have ternary N' nitrogens, as shown in Table 3.1. The N'—P bonds of 1b

and 1c are also longer than in their counterparts 3b and 3c that have a ternary N'
nitrogen. The increase in the N'—P bond length suggests that the N'—P is
weaker if the N' is quaternary.

NBO analysis shows that as expected the anomeric n(O) → σ*(N'—P)
interaction is greater in 1b than 1c with interaction energies, E(2), values of 139.1
kcal /mol and 70.8 kcal/mol, respectively for the B3LYP/6-311++G(d,p) optimized
structures. For the MP2/6-311++G(d,p) optimized structures, this value is 137.4
kcal/mol and 76.9 kcal/mol, respectively. Because the values of the anomeric
n(O) → σ*(N'—P) interaction are well-correlated to N'—P bond lengths, we
expect that the E(2) values for the same interaction to be even larger for
structure 1a. To test this and because we were unable to find a minimum
structure for 1a at any theory level tested, we performed an NBO analysis on
minimized structures for 1b with the extra proton on the phosphoryl removed.
While not a true stationary point, this calculation provides an idea of the
magnitude of the anomeric n(O) → σ*(N'—P) interaction. We have found for both
the MP2/6-311++G(d,p) and B3LYP/6-311++G(d,p) structures, the anomeric
n(O) → σ*(N'—P) interactions are increased for this restrained structure, 1a, with
values of 167.0 kcal/mol for both structures explaining its instability.

Dative bonds formed from a nitrogen lone pair as a donor are well-
documented in the literature [118,119], as well as those with phosphorus as an
acceptor [1]. Our calculations show that this dative interaction, seen by n(N) →
n*(P) energies, may provide an explanation of the dissociation of 1a. In
structures with longer N'—P bonds, lesser dative n(N) → n*(P) interactions are
expected. The NBO analysis supports such an interpretation where the n(N) →
n*(P) interactions increase with shorter N'—P bond lengths. For example, the
E(2) values for the n(N) → n*(P) interactions at the B3LYP/6-311++G(d,p) level
increases from 173.7 kcal/mol to 203.2 kcal/mol to 419.4 kcal/mol from 1a to 1b
to 1c. At the MP2/6-311++G(d,p) level these values are 173.7 kcal/mol, 216.0
kcal/mol and 411.6 kcal/mol. Taken together (Figure 11) the dissociation of the
N'—P bond in 1a occurs when the magnitude of the n(N) → n*(P) bond
stabilizing interaction is no greater than the anomeric n(O) → σ*(N'—P)

destabilizing interactions. Bond formation will occur if the dative forces are greater than the destabilizing anomeric effect, explaining why structure 1a is not a stationary point.
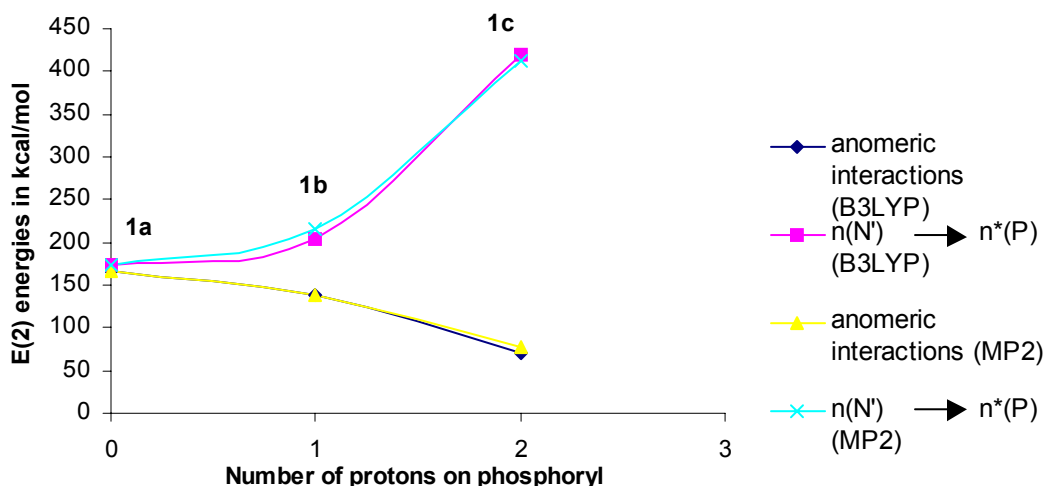
Figure 3.11. Interaction energies for series 1 structures. For structure 1a, when the N'—P bond does not form, the anomeric interactions i.e. the n(O) → n*(N'—P) interactions that favor breaking the N'—P bond are almost equal to the n(N') → n*(P) interactions that favor forming the N'—P bond. When the N'—P bond is shorter and stronger, the n(N') → n*(P) interaction (that favors forming the N'—P bond) dominates. E(2) interaction energies are used in this case.

Opposing resonance theory. The opposing resonance theory states that the weakness of the N'—P bond is due to the competition for the same lone pair on N' by two strongly delocalized groups, i.e phosphoryl and guanidinium. For opposing resonance to be a significant effect, the magnitude of the delocalization of the lone pair on N' towards the guanidinium should be approximately equal to the delocalization of the same lone pair towards the phosphoryl group. In short, the magnitude of the n(N') → σ*(C—N) and n(N') → σ*(C—N'') interactions should be nearly equal to that of the n(N') → σ*(P—O) type interactions. However, for all structures, the magnitude of n(N') interactions with C—N and C—N'' bonds are markedly greater than those with P—O bonds (Tables S8 and S9). As expected from using basic chemical intuition, this imbalance towards the guanidinium gives little competition for the lone pair and opposing resonance is not a strong effect.

To support opposing resonance theory, it is also expected that a balance of n(N') → σ*(C—N'') and n(N') → σ*(C—N) with n(N') → σ*(P—O) interactions

to be greater in structures with longer N'—P bonds. For example, the (N') → σ*(C—N) and n(N') → σ*(C—N'') interactions and the n(N') → σ*(P—O) interactions should be larger and better balanced in 3a than 3c. Again the results show that this is not the case when comparing all structures (Tables 3.6 and 3.7). For example, for structure 3a, the values are 50.6 kcal/mol for the n(N') → σ*(C—N) and n(N') → σ*(C—N'') interactions compared to 7.2 kcal/mol for the n(N') → σ*(P—O) interactions. For structure 3c, the values are 27.5 kcal/mol for the n(N') → σ*(C—N) and n(N') → σ*(C—N'') interactions compared to 23.6 kcal/mol for the n(N') → σ*(P—O) interactions. In fact, this is opposite to opposing resonance theory with the structure with the shorter N'—P bond experiencing more competition for the lone pair of N' and therefore less overall resonance than the structure with the longer N'—P bond.

In fact no relationship between the n(N') → σ*(C—N) and n(N') → σ*(C—N'') as well as n(N') → σ*(P—O) interactions and N'—P bond lengths can be derived from the interaction energies (Tables 3.7 and 3.8). The computed data underscores the inability of opposing resonance theory to describe N'—P bond lability in phosphagens.

Relevance to high-energy phosphoryl bonds. The lability of high-energy phosphoryl bonds is crucial for many cellular processes[120], in particular, cellular energy buffering by phosphagens and energy production by ATP hydrolysis [18,19]. However, detailed structural studies have been limited due to the lability of these high-energy bonds possibly from experimental resolution and time scale issues. In modern textbooks [18,19], a qualitative understanding based upon the opposing resonance effect, resonance stabilization and solvation have been used to explain this phenomenon. For the first time, we have shown that the anomeric effect plays a significant role in the weakness of the N'—P bond and not the

Table 3.7 : Interactions of n(N') with other orbitals on B3LYP/6-311++G(d,p) optimized structures.

Donor	Acceptor	E(2) kcal/mol	E(j)-E(i) a.u.	F(i,j) a.u.
Structure **3a**				
n(N')	σ*(C-N)	0.82	1.06	0.028
n(N')	σ*(C- N")	20.78	1.02	0.134
n(N')	σ*(C-N")	39.20	0.68	0.148
n(N')	σ*(P-O1)	3.90	1.03	0.059
n(N')	σ*(P-O2)	3.00	1.03	0.051
Structure **3b**				
n(N')	σ*(C-N)	0.63	1.13	0.025
n(N')	σ*(C-N")	10.52	1.15	0.101
n(N')	σ*(C-N")	32.58	0.71	0.138
n(N')	σ*(P-O1)	1.53	1.05	0.036
n(N')	σ*(P-O2)	12.99	0.84	0.093
n(N')	σ*(P-O3)	1.95	1.04	0.041
Structure **3c**				
n(N')	σ*(C-N")	3.39	1.32	0.061
n(N')	σ*(C-N")	31.32	0.69	0.134
n(N')	σ*(P-O1)	5.52	0.87	0.062
n(N')	σ*(P-O2)	17.75	0.89	0.112
n(N')	σ*(P-O3)	1.84	1.06	0.04
Structure **2b**				
n(N')	σ*(C-N)	128.9	0.47	0.228
n(N')	σ*(C-N")	7.02	1.14	0.085
n(N')	σ*(P-O1)	6.02	0.87	0.067
n(N')	σ*(P-O2)	2.16	1.07	0.045
Structure **2c**				
n(N')	σ*(C-N)	83.23	0.53	0.197
n(N')	σ*(P-O1)	8.89	0.9	0.081
n(N')	σ*(P-O2)	6.27	0.91	0.069

Table 3.8: Interactions of n(N') with other orbitals on the MP2/6-311++G(d,p) optimized structures.

Donor	Acceptor	E(2) kcal/mol	E(j)-E(i) a.u.	F(i,j) a.u.
Structure **3a**				
n(N')	σ*(C-N)	1.32	1.10	0.035
n(N')	σ*(C-N")	18.85	1.12	0.134
n(N')	σ*(C-N")	30.42	0.69	0.131
n(N')	σ*(P-O1)	2.00	1.07	0.043
n(N')	σ*(P-O2)	1.57	1.05	0.037
n(N')	σ*(P-O3)	3.65	1.08	0.058
Structure **3b**				
n(N')	σ*(C-N)	1.04	1.15	0.032
n(N')	σ*(C-N")	12.63	1.18	0.111
n(N')	σ*(C-N")	24.81	0.73	0.122
n(N')	σ*(P-O1)	0.80	1.08	0.027
n(N')	σ*(P-O2)	11.56	0.87	0.090
n(N')	σ*(P-O3)	2.42	1.07	0.046
Structure **3c**				
n(N')	σ*(C-N)	0.83	1.21	0.029
n(N')	σ*(C-N")	8.06	1.26	0.092
n(N')	σ*(C-N")	18.59	0.76	0.108
n(N')	σ*(P-O1)	4.52	0.9	0.057
n(N')	σ*(P-O2)	17.19	0.92	0.112
n(N')	σ*(P-O3)	1.91	1.09	0.042
Structure **2b**				
n(N')	σ*(C-N)	130.04	0.48	0.229
n(N')	σ*(C-N')	11.34	1.13	0.108
n(N')	σ*(P-O1)	5.17	0.87	0.062
n(N')	σ*(P-O2)	2.75	1.07	0.051
Structure **2c**				
n(N')	σ*(C-N)	82.03	0.54	0.197
n(N')	σ*(P-O1)	9.03	0.91	0.083
n(N')	σ*(P-O2)	5.96	0.92	0.067

opposing resonance effect. The anomeric effect also explains the trend that increasing phosphoryl resonance leads to longer N'—P bonds. The symmetry between the phosphate groups in ATP could imply a more balanced competition for the lone pair on the bridging oxygen atom, giving a more significant opposing resonance effect. Nevertheless, the computed evidence of the significant role the anomeric effect plays in weakening the N'—P bond in phosphagens requires such reinforcing anomeric effects to be considered in the lability of the O—P bond in ATP. In particular, the anomeric control of high-energy bonds is likely to

impact the understanding of phosphoryl transfer reaction mechanism and catalysis.

Conclusion

Modulating the resonance capability of N-methyl-N'-phosphorylguanidines by studying structures differing in both the number and placement of protons provides a novel interpretation of the stereoelectronic factors contributing to the lability of the high-energy N'—P bond in phosphagens. Compelling evidence shows that decreased delocalization at either the phosphoryl or guanidinium groups of N-methyl-N'-phosphorylguanidine results in stronger N'—P bonds in accordance with the traditional view that the phosphoryl and guanidinium are more strongly stabilized by resonance when separated than when bonded. In particular, an explicit link between different protonation states and N'—P bond lengths is provided by the quantum computations. For example, protonation at N'' alone is sufficient to cause significant weakening of the N'—P bond. However, evidence of competition between the phosphoryl and guanidinium groups for the same lone pair on the bridging nitrogen, N', is not found as described in opposing resonance theory. Indeed, we find that interactions between the lone pair on the bridging nitrogen, N', and the guanidinium dominates over interactions between the same lone pair and the phosphoryl thus negating the possibility of competition between the two groups. The weakness of the N'—P bond is directly correlated with an increase in the anomeric interaction n (O) $\rightarrow$ σ^*(N'—P). To a lesser extent, interactions of σ(N'—P) $\rightarrow$ σ^*(C—N) and σ(N'—P) $\rightarrow$ σ^*(C—N'') type can also explain the weakness of the N'—P bond. These interactions directly correlate N'—P bond strength with protonation state and provide a rationalization for an increase in overall resonance being accompanied by a weakening of the N'—P bond. The anomeric effect in particular provides an intriguing new avenue into the factors controlling the strength of high-energy bonds and phosphoryl transfer reactions.

ON THE NATURE OF THE HIGH ENERGY BOND

<u>Introduction</u>

Previous work identified the generalized anomeric effect as a factor for the weakness of high energy N—P phosphoester bonds in N-phosphorylguanidines [121]. The generalized anomeric effect is a delocalization effect where electron density is delocalized from the antiperiplanar lone pairs from the three phosphoryl oxygen atoms into the single O—P antibonding orbital, thus weakening that bond [24,25,122]. In the NBO nomenclature, this is seen as three independent $n(O) \rightarrow \sigma^*(O—P)$ interactions. The generalized anomeric effect stabilizes a double bond/no bond resonance form [25]. In the case of phosphate compounds, such as N-phosphorylguanidines, the no bond stabilized form is conveniently the same high energy bond cleaved by hydrolysis. Physically, this phenomenon is manifested by a lengthening of the high energy bond. Energetically, the same phenomenon results in an energy lowering of the products, when the bond in question is indeed absent, giving a more exothermic free energy of hydrolysis. Here, we show that the generalized anomeric effect is not just an isolated effect accounting for the weakness in phosphoguanidinium compounds, but it rather deeply impacts biology by controlling the weakness of the cleaved O—P bond in important phosphoryl compounds such as ATP. First, we establish a link between the weakness (bond length) of the cleaved O—P phosphoryl bond and experimental standard free energies of hydrolysis for a series of phosphoryl compounds, Figure 4.1, ranging in phosphoryl transfer potential. Second, we show that both O—P phosphoryl bond lengths and experimental standard free energies of hydrolysis correlate strongly with the magnitude of the bond weakening anomeric effect, as calculated by natural bond orbital (NBO) analysis [42]. Third, we investigate how different environmental effects affect bond strengths and show that the magnitude of anomeric interactions and consequently bond

lengths is modulated by the presence of solvent, protonation and coordination to Mg^{2+}.

$$X—O\sim\sim\sim P—O^-$$

with O^- above P and O (double bonded) below P

cleaved O-P phosphoryl bond
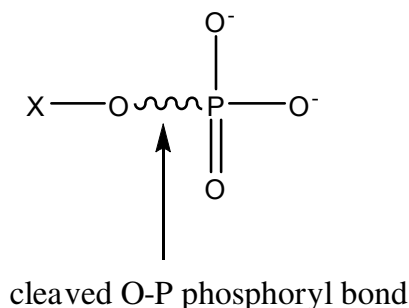
Substituent	Experimental dG hydrolysis (kcal/mol)
X= 1: glycerate (glycerol-3-phosphate)	-2.2
2: methanolate (methyl monophosphate)	-2.2
3: 3,4,5,6 tetrahydroxytetrahydropyranyl methanolate (glucose-6-phosphate)	-3.3
4: 3,4,5 trihydroxy 6 tetrahydropyranolate (glucose-1-phosphate)	-5.0
5: pyruvate enol (phosphoenol pyruvate)	-6.8
6: methyl monophosphate (methyl diphosphate)	-8.5
7: methyl diphosphate (methyl triphosphate)	-8.5
8: acetate (acetyl phosphate)	-10.3
9: 3-phosphoglycerate (1,3 bisphosphoglycerate)	-11.8
10: carbamate (carbamyl phosphate)	-12.5

Figure 4.1 : A schematic of the structures studied in this work. Different substituents are described and standard free energies of hydrolysis given. Note: for phosphoenol pyruvate only the free energy of the phosphoryl transfer process of hydrolysis is given. The tautomerization from the enol to keto form is not included. Names of final structures are shown in parentheses.

Methods

All electronic structure calculations were carried out using G03 [76]. Natural bond orbital calculations were performed using NBO 3.1 [42] as implemented in G03. Initial conformations for all structures were obtained from either prior published

calculations [123,124], crystal structures [125,126] or potential energy scans about select dihedrals at the PM3 [80] semiempirical molecular orbital approximation. Minimum energy structures from these were optimized using B3LYP [84,85] density functional method and the 6-311++G(d,p) [98,99] basis set. All calculations for structures in Figure 1 were performed in water using the PCM [127] model. Calculations of the magnitude of the anomeric effect were performed using NBO 3.1[42]. NBO transforms the non-orthogonal atomic orbitals from the HF wavefunction into natural atomic orbitals (NAO), natural hybrid orbitals (NHO) and natural bond orbitals (NBO) each of which are complete and orthonormal. This allows electron density to be localized onto bonds and atoms, leading to a better description of the molecule as a localized Lewis structure. Because NBO transformation provides filled orbitals that are more concentrated (localized) in terms of occupancies delocalizing interactions are then effectively treated as a perturbation through second-order perturbation theory. The E(2) energy values from the second-order perturbation method then provide a reasonable quantitative description of the magnitude of such delocalizing interactions [22,122]. Another way of obtaining a quantitative estimate of the magnitude of delocalizing interactions is the deletion method as implemented in NBO. Here, the interacting orbitals are removed from the Fock matrix in NBO and the difference in energy between the two density matrices, with and without the interacting orbitals is calculated. This is called the deleted energy. Because the deletion method can assess the net effect of more than 1 interaction, it can be used to assess the possibility of anti-cooperative orbital interactions which E(2) values alone do not reflect [122]. NBO deletion analysis was performed using the HF/6-311++G(d,p) method on B3LYP/6-311++G(d,p) optimized structures due to the inconsistency of deletion operations using DFT wavefunctions, as detailed in the NBO manual [42]. For the study of environmental effects on triphosphate structures, only gas phase calculations were used as the intention was to probe the effect of protonation and electrostatics only. Prior calculations on the *high energy* N—P bond in phosphagens [121] and tests on selected structures showed that this level of theory was sufficient to identify the changes in phosphoryl bond lengths associated with the generalized anomeric effect. More extensive basis sets and

higher levels of theory did not change any general trend. Frequency calculations were performed in conjunction with all energy minimizations to characterize all stationary points as minima and not transition structures. Charges were calculated according to the Merz-Singh-Kollman method [128,129].

<u>Results and Discussion</u>

The phosphoryl compounds modeled, as shown in Figure 4.1, were chosen to span the wide range of free energies of hydrolysis from high energy (-12.5 kcal/mol) to low energy (-2.2 kcal/mol) systems. Optimized structures are shown in Figure 4.2.

Hereafter, the cleaved high energy bond as shown in Figure 4.1, is referred to simply as the O—P bond. A direct correlation (R=-0.89) between experimental standard free energies of hydrolysis and calculated O—P bond lengths is found, as shown in Figure 4.3a. This finding clearly shows that not all O—P bonds are equal in strength. The correlation between O—P bond lengths and free energies of hydrolysis strongly suggests that bond weakening effects contribute to exothermic free energies of hydrolysis.

The observed correlation is found to stem from stereoelectronic effects inherent in the phosphoryl group. Specifically, O—P bond lengths are found to correlate (R = 0.93) with the n(O)→σ*(O—P) generalized anomeric effect (Figure 4.3b) using NBO [42]. The NBO method describes the molecule as a Lewis structure with electron density localized onto bonds and atoms [109]. Delocalizing interactions are then quantitatively described as perturbations to this localized system using perturbation theory. The anomeric energy also correlates with the experimental standard free energies of hydrolysis (R = -0.95) (Figure 4.3c), further underlining the connection between high energy status and the magnitude of the anomeric effect. To our knowledge, no other molecular reason given for high energy nature [3,12,14,69,130,131] correlates with this range of experimental free energies of hydrolysis.
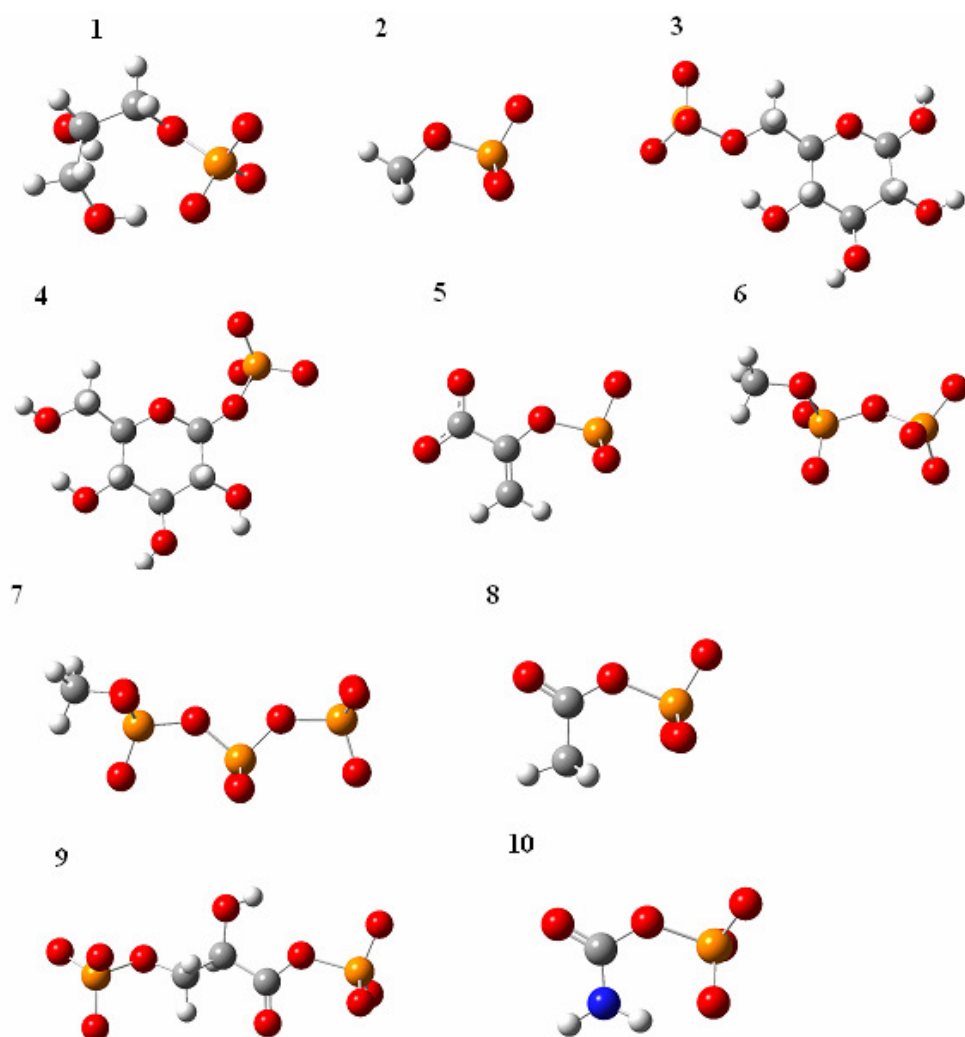
Figure 4.2 Optimized structures
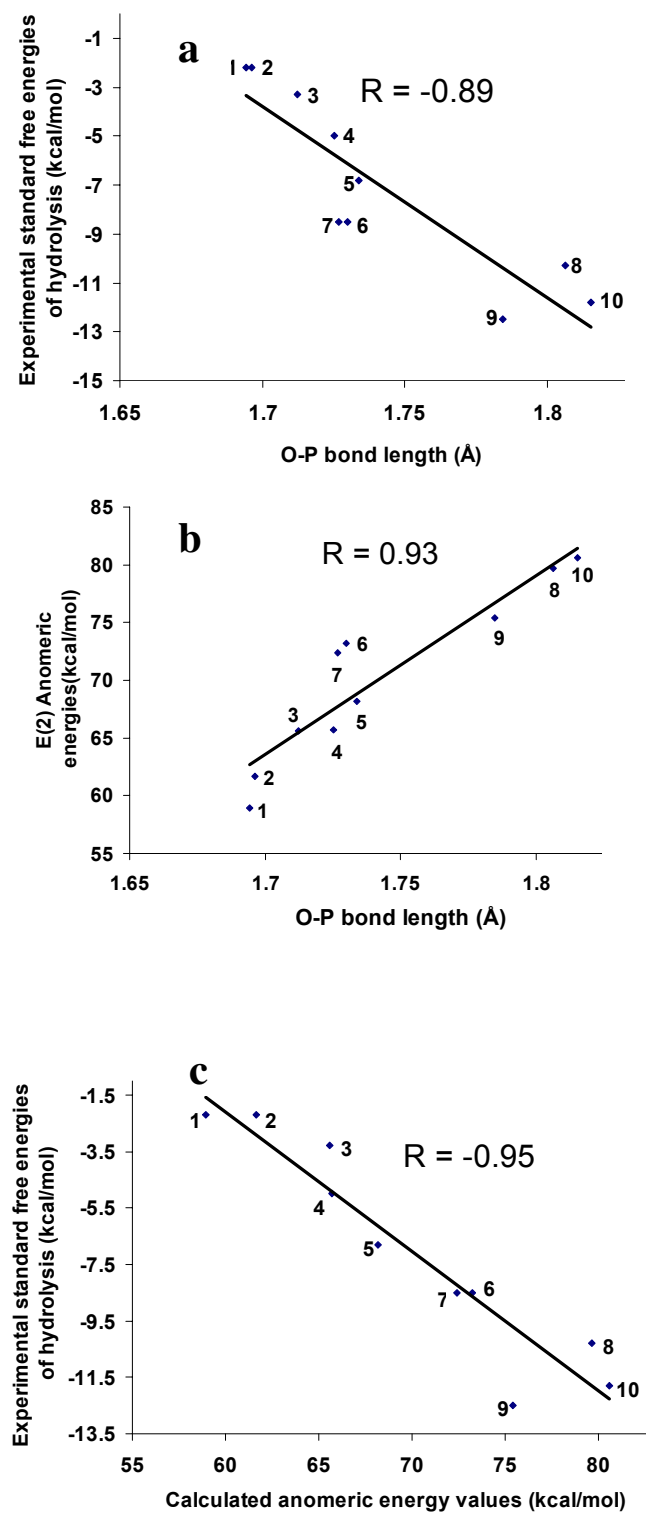
Figure 4.3 a) Correlation between standard free energies of hydrolysis and O—P bond lengths. b) Correlation between O—P bond lengths and the magnitude of the $n(O){\rightarrow}\sigma^*(O{-}P)$ anomeric effect. All structures were calculated at the B3LYP/6-311++G(d,p) level of theory in water using the PCM method. c) Correlation between standard free energies of hydrolysis and the magnitude of the $n(O){\rightarrow}\sigma^*(O{-}P)$ anomeric effect.

The n(O)→σ*(O—P) generalized anomeric effect is a delocalization or stereoelectronic effect, whereby electron density from the antiperiplanar lone pair of electrons is delocalized into an adjacent antibonding orbital on polar bonds [22,24,25,43,122]. This is also known as a hyperconjugative or charge transfer effect [25]. The O—P—O [43] moiety along the phosphoryl compound ideally satisfies all conditions needed for the anomeric effect to occur. The O—P bond is polar and each oxygen atom on the phosphoryl group has one antiperiplanar lone pair (for a total of three such interactions) available for the donation of electron density into the antibonding σ*(O—P) orbital. As expected, the results show that only antiperiplanar lone pairs participate in anomeric delocalization. The magnitude of each anomeric interaction is ~25 kcal/mol, which is typical of other reported charge transfer interactions [43,132]. The net anomeric effect is a sum of contributions from all three oxygen atoms. No evidence of anti-cooperative effects, for example destructive interference between the three contributions is seen, Figure 4.4. As detailed in previous work [122] an analysis of the combined effect of electron density donations to the same antibonding orbital is important when investigating any general trend. This is because anti-cooperative effects i.e. destructive interference between lone pair wavefunctions can remove any correlation. To check that no anti-cooperative effects occur, the deletion method as implemented in NBO 3.1 is used. Here the combined effect of eliminating all 3 interactions is compared to the effect of individually deleting each interaction and numerically adding all three energy terms. If the combined elimination is less than the numerical addition, then anti-cooperative effects occur. The convention followed here is that anomeric energies refer to E(2) values, calculated using density functional theory, from the three phosphoryl oxygen atoms numerically summed.
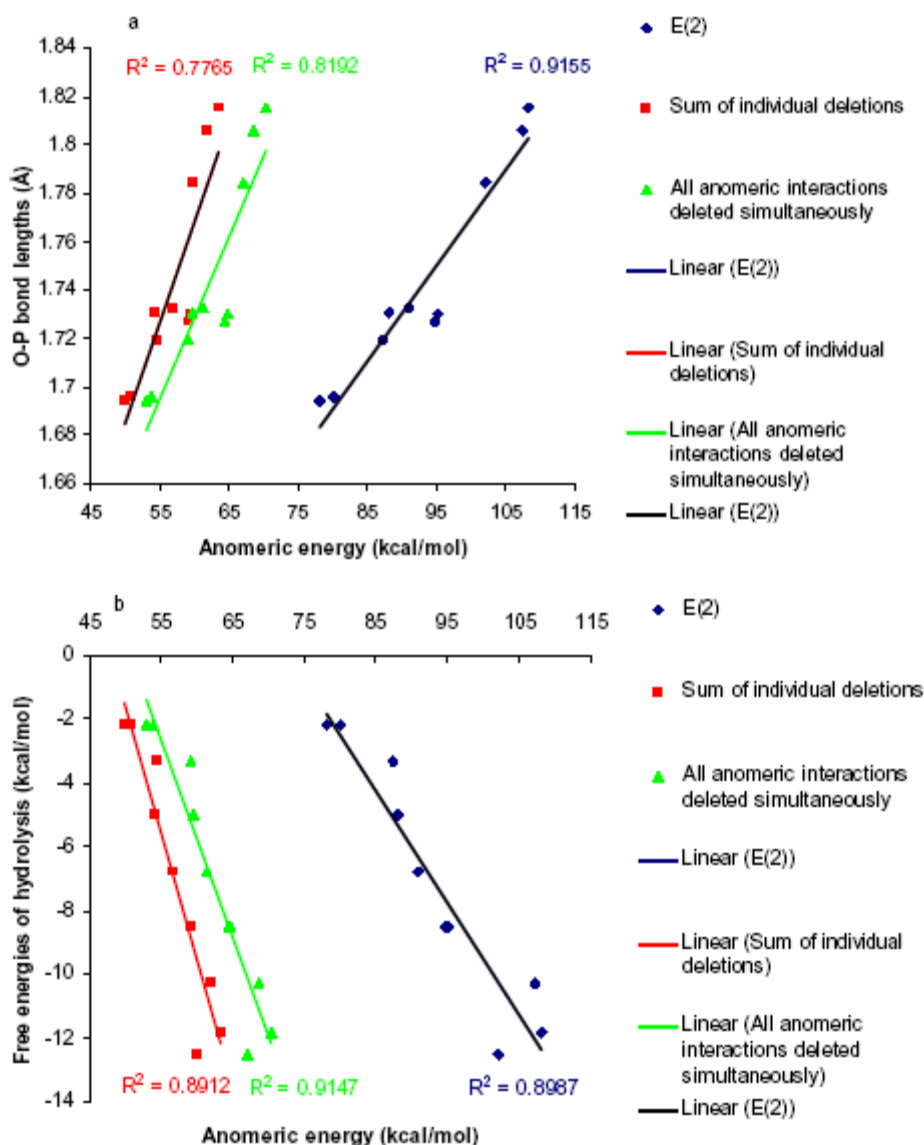
Figure 4.4 a Correlation between O—P bond lengths and anomeric energy. **b** Correlation between free energies of hydrolysis and anomeric energy. In both cases the anomeric energy was calculated using three different methods. 1) second order perturbation (E(2)) 2) the deletion procedure by summing individual deletions and 3) the deletion procedure where all interactions were deleted simultaneously.
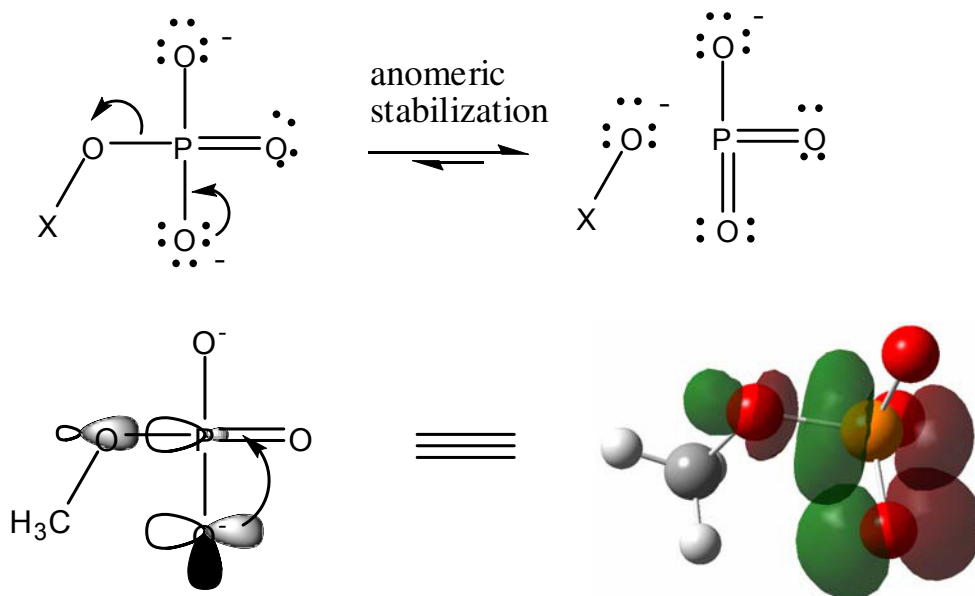
Figure 4.5 A schematic of the anomeric effect in phosphates

To understand how the anomeric effect impacts the exothermicity of hydrolysis, it is instructive to consider a simple resonance picture, as shown in Figure 4.5. The anomeric effect preferentially stabilizes a double-bond/no-bond resonance structure relative to the bonded form. Conveniently, the no-bond preferred is also the exact O—P bond broken during the hydrolysis process. Thus, the resonance form stabilized by the anomeric effect is essentially the products of the reaction, and an enhanced anomeric effect results in products of the reaction that are lower in energy or a more exothermic hydrolysis. The relative destabilization of the reactants is physically manifested by a distinct weakening of the O—P bond, by electron density donation into σ*(O—P), corroborating the intuitive idea where a weaker O—P bond leads to a less endothermic bond dissociation process allowing for a more exothermic overall free energy of hydrolysis.

Product stabilization by the anomeric effect is distinct from other resonance explanations given in biochemical texts [3,12,69,130]. These focus on the increased resonance stability of the substituent ("X" in Figure 4.1) relative to the phosphorylated form. Such explanations involve delocalization effects along the

substituent group where the number of alternate resonance structures for substituent X is greater when not phosphorylated. The anomeric effect is different from these explanations, as it does not involve any delocalization interactions on the substituent group. The anomeric effect is solely dependent upon the magnitude of interaction between the lone pairs on the phosphoryl and σ^*(O—P) antibonding orbital

Charges calculated according to the Merz-Singh-Kollman method [128,129] show that in general, negative charges on the phosphoryl oxygens decrease as the anomeric effect increased. Although the decrease in charges were small ~0.0014 e, the trend nevertheless matches with the expectation that as electron density from the phosphoryl oxygens is donated into the σ^*(O—P) antibond the less negative the overall charge.

*Enhancement of the anomeric effect in high energy compounds by substituent effects.*The magnitude of the anomeric effect is often described as being dependent upon the energy gap between the lone pairs and antibonding orbitals [22,122]. To understand why the anomeric effect is enhanced in high energy compounds as opposed to low energy compounds, we investigated the orbital energies of both the lone pair and O—P antibond. The dependency of hydrolysis free energies and the anomeric effect upon orbital energy gaps is clearly evident in the computed orbital energies. As shown in Figure 4.6a, compounds with larger hydrolysis energies show a smaller gap between lone pair and antibonding orbital energies. The finding is underscored in Figure 4.6b, where a greater anomeric effect is given by a smaller gap between orbital energies. It is also evident from Figure 4.6 that the closing of the gap is primarily due to the lowering of the σ^*(O—P) orbital energy, rather than the raising of the lone pair orbital energy.
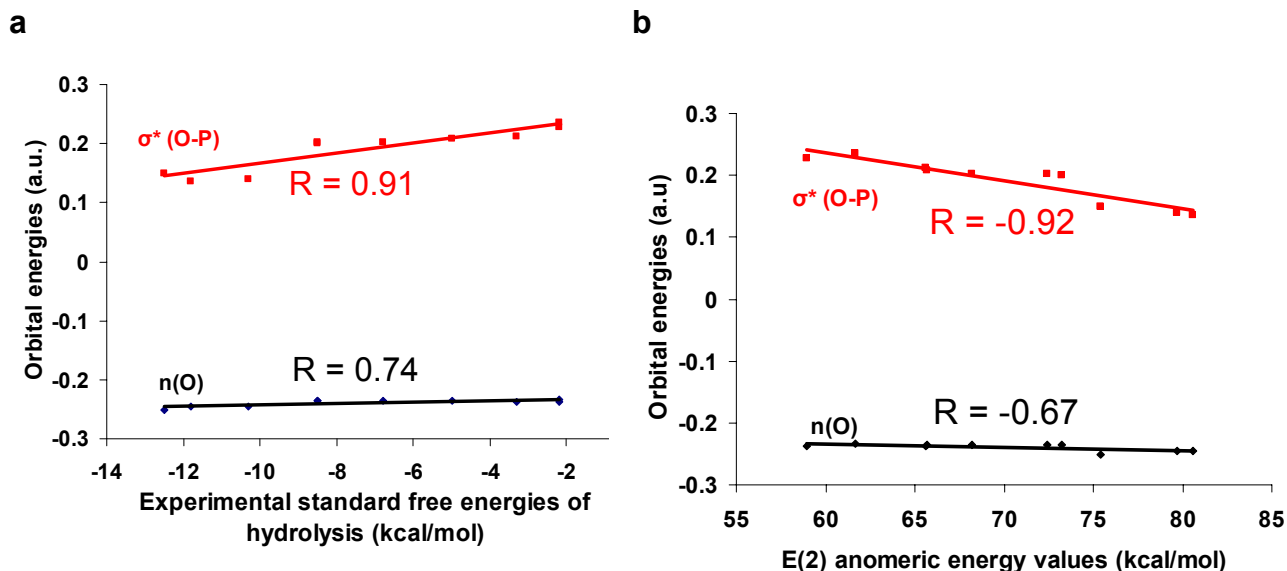
Figure 4.6. a) Relationship between orbital energies and standard free energies of hydrolysis. b) Relationship between orbital energies and anomeric energies.

A relationship between σ* orbital energies and polarity is found, as shown in Figure 4.7. Using the difference between the percentages of NBO on O and P as an index of polarity (this is also equal to the difference between squares of polar coefficients) of the O—P bond, an increased polarity of the σ*(O—P) bond is found to correlate with a lower σ*(O—P) orbital energy, as shown in Figure 4.6a. With a lower σ*(O—P) orbital energy, the energy gap between the lone pairs and σ*(O—P) is reduced leading to an enhanced anomeric effect and in turn more exothermic free energies of hydrolysis, as shown in Figures 4.7b and 4.7c. Others have also found that the lowering in energy of orbitals is often due to an increase in polarity [22]. The substituent enhances the polarity of the σ*(O—P) antibonding orbital by inductively withdrawing electrons. In general, it is found that the greater the electron withdrawing capacity of the substituent group (C=O and COR in acetate and 3-phosphoglycerate), the greater the polarity of the σ*(O—P) antibonding orbital and the greater the magnitude of the anomeric effect.

a

b

c
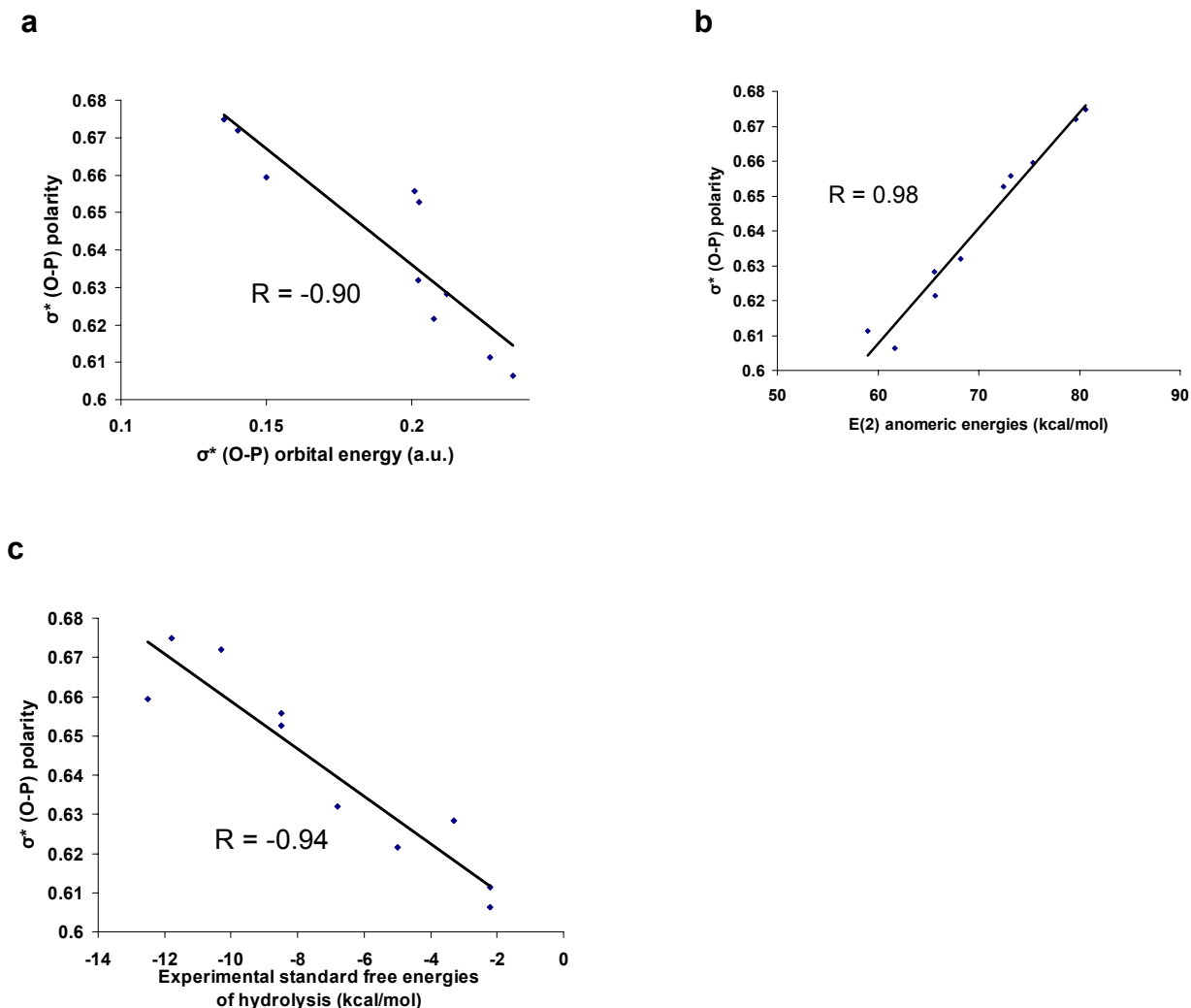
Figure 4.7. a) Correlation between σ* (O—P) polarity and orbital energies. b) Correlation between σ* (O—P) polarity and anomeric energies. c) Correlation between σ* (O—P) polarity and hydrolysis energies.

Substituent effects upon other resonance factors. A common understanding of *high energy* character stemming from Kalckar's opposing resonance theory is that the electron-withdrawing substituent removes electron density from the bonding orbital leading to weaker O—P bonds. We investigated electron delocalization from the bonding, σ(O—P), Figure 4.8, using NBO. A correlation is found between the removal of electron delocalization away from the σ(O—P) bond with longer O—P bond lengths and greater experimental free energies of hydrolysis. However, the correlation is not as high as it is with the anomeric effect. More importantly, NBO shows that the magnitude of these interactions is significantly less than the anomeric effect. E(2) values for these

delocalizations hardly reach 10 kcal/mol whilst even one lone pair anomeric interaction reach values of ~ 20 kcal/mol.

a   b
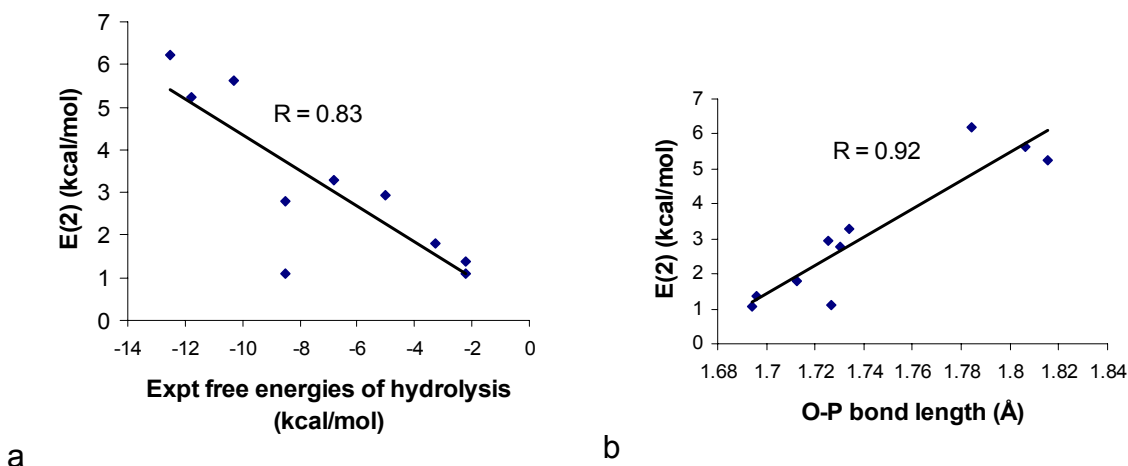
Figure 4.8. a) Correlation between σ(O—P) delocalizations into the substituent group and experimental free energies of hydrolysis. b) Correlation between σ(O—P) delocalizations into the substituent group and O—P bond length.

The comparative lack of delocalization away from the σ(O—P) as compared to the delocalization into the σ*(O—P) becomes even more important when orbital occupancies are examined in Figure 4.9, where it is shown that σ*(O—P) occupancy increases with more exothermic free energies of hydrolysis and increased O—P bond lengths. In contrast, not only does the σ(O—P) have a weaker correlation with bond length and hydrolysis energies, but its expected decrease in occupancy with more exothermic free energies of hydrolysis and increased O—P bond lengths is not as apparent as the increase in σ*(O—P) occupancy.
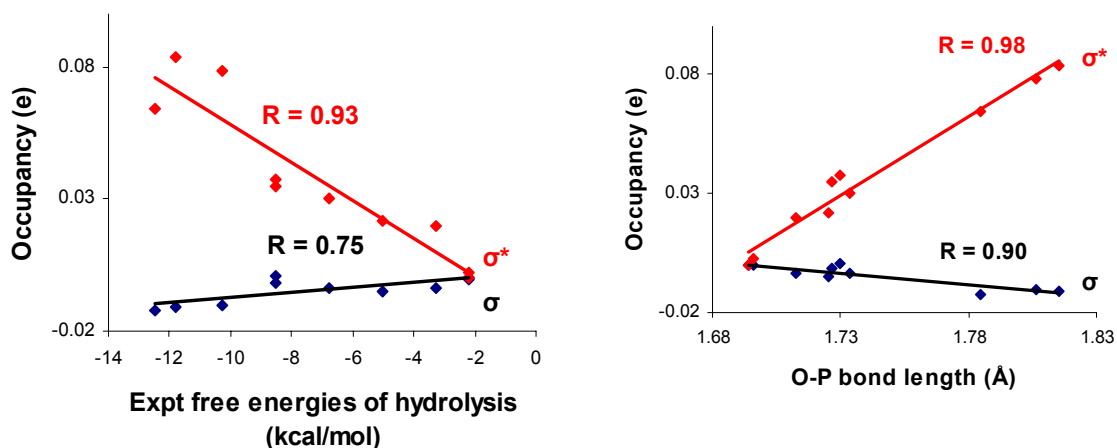
Figure 4.9. σ(O—P) and σ*(O—P) occupancies and their correlation with experimental free energies of hydrolysis. σ(O—P) and σ*(O—P) occupancies and their correlation with O—P bond length.

The NBO results imply that compared to the delocalization away from the bonding orbital, the anomeric effect may be a more important effect in O—P bond weakness. Delocalization from a bonding orbital induces weakness in that bond. Delocalization into an anti-bonding orbital also induces weakness in that bond. The occupancy of a bonding orbital is usually much greater than an anti-bonding orbital. Electron withdrawing effects of the substituent has an overall weaker effect on overall orbital occupancy of a nearly-filled orbital than electron donation into a largely empty antibonding orbital. The electron withdrawing effect of the substituent even further empties the antibonding orbital allowing for even greater delocalization from the lone pairs. The NBO results fit in nicely to the traditional idea that lone pairs are much better donors than σ bonds. Hence, while electron withdrawing effects of the substituent upon the bonding orbital will certainly have some effect on the weakness of the O—P bond and the traditional physical organic picture of substituent effects cannot be disregarded, neither can the anomeric effect.

Environmental effects. Environmental factors that change lone pair donor and antibonding acceptor abilities affect the magnitude of the anomeric effect and subsequently O—P bond lengths and stability. The effect of the environment on O—P bonds is perhaps most clearly evident in a series of FTIR experiments on phosphate monoesters in different solvents performed by Cheng and coworkers [133]. Polar solvents such as water are known to reduce the overall anomeric effect by sequestering lone pairs into hydrogen bonding networks thus reducing the overall donation into antibonding orbitals. Cheng's experiments show a distinct lengthening of the bridging O—P bond in less polar solvents such as DMSO compared to water. This experiment was modeled here with a simple methyl monophosphate using the B3LYP/6-311++G(d,p) level of theory. Methyl monophosphate, coordinated with three water molecules and optimized in water using PCM [127,134] was compared to methyl monophosphate coordinated to three DMSO molecules optimized in DMSO also using PCM. It was found that the reactive O—P bond for a simple methyl monophosphate did increase ~ 0.03 Å (experimental value ~0.015 Å) in moving from water to DMSO. This increase in bond lengths was matched by an increase in E(2) values for the anomeric effect by ~11 kcal/mol showing once again the connection between O—P bond lengthening and the anomeric effect.

To further investigate how environmental factors influence the anomeric effect, we performed electronic structure calculations on five methyl triphosphate molecules (a model of ATP). Three of these structures were chosen to increase the protonation state sequentially at the terminal γ-phosphoryl group: 11 – deprotonated ; 12 - protonated once ; 13 - protonated twice (Figure 7). The aim of choosing three protonation states was to reduce sequentially the anomeric effect within the terminal phosphoryl and observe the effects upon the geometry of the O—P bond. Protonation reduces the anomeric effect by removing electron lone pair donation into the Os—Pγ antibonding orbital.

The remaining two structures compare two different Mg^{2+} chelated schemes in order to observe the effects of electrostatics upon Os—Pγ bond lengths. Structure 14 is coordinated with oxygen atoms on Pα, Pβ and Pγ. This

type of coordination is slightly unusual but has been reported for the transition state analog of arginine kinase [17]. A more common coordination scheme is shown in structure 15, where Mg^{2+} is coordinated with oxygen atoms on Pα and Pβ [135]. Three and four water molecules complete the coordination of Mg^{2+} in systems 14 and 15, respectively. These structures were chosen to establish the impact of electrostatics modulating the anomeric effect on phosphoanhydride bond length changes.
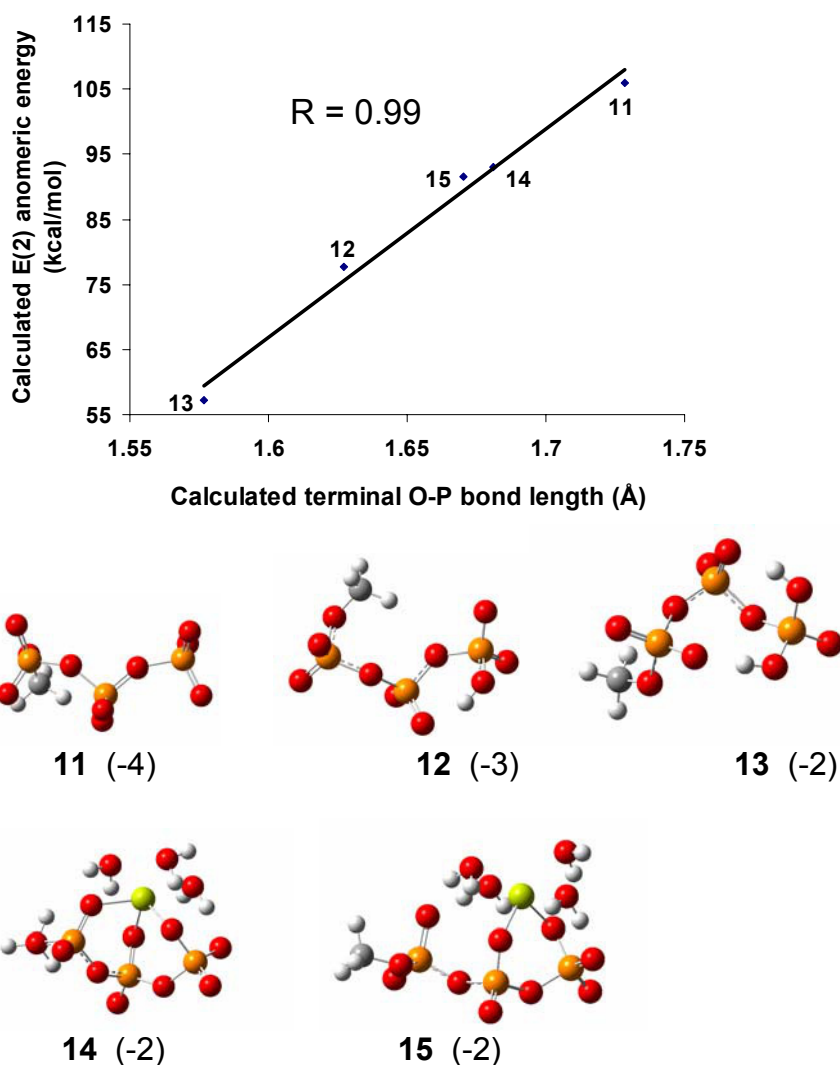
Figure 4.10. a) Correlation between terminal O—P bond lengths and the magnitude of the n(O)→σ*(Os—P) and optimized geometries of triphosphate structures

As shown in Figure 4.10, there is a clear geometric shortening of the terminal O—P bond with increasing protonation or coordination of Mg^{2+}. The computed results emphasize that the geometric changes within triphosphates are

58

caused by the anomeric effect. Charge transfer between Mg^{2+} and the lone pairs on oxygen atoms on the phosphoryl groups reduces the availability of the same lone pairs to donate into the antibonding $\sigma^*(O-P)$ orbitals, thus strengthening the O—P bond. The shortening of the terminal O—P bond with increasing protonation reiterates how reducing lone pair donor ability strengthens the O—P bond.

Energetic consequence of the anomeric effect. The correlation shown between the O—P bond weakening anomeric effect and experimental free energies of hydrolysis underlines how the anomeric effect mediates the high energy nature of phosphoryl compounds. A paired t-test between standard free energies of hydrolysis and the magnitude of the $n(O)\rightarrow\sigma^*(O-P)$ anomeric effect gives a P-value of 3.0×10^{-10} indicating that there is little chance that anomeric effect is correlated to phosphoryl transfer potential by coincidence.

As mentioned by Kirby, stereoelectronic effects such as the anomeric effect are easily observed, but whose consequences are more difficult to measure energetically. [25] In order to approximate the importance anomeric contribution to the standard free energies of hydrolysis, protonation was used as a means to remove the anomeric effect and gauge its consequences. The anomeric effect is eliminated by protonation because the antiperiplanar lone pairs are removed from the stereoelectronic interaction. We considered two similar monophosphates, acetyl phosphate and methyl monophosphate. Acetyl phosphate is high energy with a standard free energy of hydrolysis of -10.5 kcal/mol. Methyl monophosphate is low energy with a free energy of hydrolysis of -2.2 kcal/mol. The difference in O—P bond lengths between the two compounds is ~ 0.11 Å. Double protonation of both compounds at the phosphoryl group removes most of the anomeric effect resulting in similar O—P bond lengths differing by only 0.03 Å. Comparing the difference in protonation energy between the two compounds provides a reasonable estimation of how much the O—P bond weakening anomeric effect is enhanced in acetyl phosphate over methyl monophosphate is obtained. The difference in protonation energy is an accurate assessment of the anomeric effect, as both compounds are structurally similar

enough that other consequences of protonation, such as steric effects and conformational changes essentially cancel. We find that the anomeric effect is enhanced in acetyl phosphate by 9.6 kcal/mol or about the same as the difference in free energies of hydrolysis, 8.3 kcal/mol, between the two compounds. This value is close to the values given by NBO. Additionally, O—P bond destabilization has been experimentally observed [133,136]. Estimations of O—P bond destabilization using vibrational frequencies to approximate force constants and bond energy result with values ranging from ~1 kcal/mol per 0.1 Å [133] stretch to ~ 15 kcal/mol per 0.1 Å stretch [136]. Our estimations of 8.7 kcal/mol per 0.1 Å stretch falls between the values from the two independent studies.

Experimental studies show that increased pH leads to increases in the exothermicity of hydrolysis in phosphates [137]. Other effects cannot be discounted from playing a role in discriminating phosphoryl transfer potential differences. However, the correlations derived in this work, together with our estimations of the contribution of the anomeric effect, and independent experimental evidence [9,137] are significant enough to underline a relationship between free energies of hydrolysis and the anomeric effect. All evidence at hand strongly suggests that the anomeric effect is important to the high energy nature of biologically important phosphoryl compounds.

Biological implications. The anomeric effect is a unique property of phosphoryl compounds shown to be significant in discriminating phosphoryl transfer potentials by alternate strengthening or weakening of reactive O—P bonds. Enzymes could potentially exploit this property to achieve catalysis. Another important characteristic of phosphates in nature is their kinetic stability in solution [4,5]. Low kinetic rates make the compounds suitable for energy storage and transport which can be overcome by enzyme catalysis. This highlights the importance of environment to reaction rates and catalysis on phosphoryl transfer reactions. Understanding the sensitivity of phosphate ester and anhydride compounds and in particular, the reactive O—P bond, to the local environment is therefore of paramount importance in understanding the mechanism of these enzymes. As demonstrated, the generalized anomeric effect offers a convenient

means for phosphoryl compounds to be sensitive to the environment - the strength of the reactive O—P bond could, in principle, be modified by the environment in an active site, enabling enzyme controlled use of otherwise stable phosphoryl compounds.

<u>Conclusion</u>

Compelling evidence is presented suggesting the role of the O—P bond weakening anomeric effect in discriminating phosphoryl transfer potentials in a range of biologically important phosphoryl compounds. Strong correlations between phosphoryl transfer potentials, O—P bond weakening, and the magnitude of the $n(O) \rightarrow \sigma^*(O—P)$ anomeric effect has been shown. Higher energy compounds show distinctly longer O—P bonds indicating that weakness of the cleaved phosphoryl bond is important in raising the overall exothermicity of the hydrolysis process by reducing the endothermicity of the bond breaking process. The anomeric effect is magnified in high energy compounds by a closer energy gap between $n(O)$ and $\sigma^*(O—P)$ orbitals. It is found that the smaller energy gap in higher energy compounds is primarily due to the lowering in energy of the $\sigma^*(O—P)$ orbital by electron withdrawing substituents that polarize that orbital and so enhance the anomeric effect. The generalized anomeric effect provides an intriguing new interpretation on how phosphoryl compounds are sensitive to the environment. Protonation, electrostatics and solvent are shown to alter the magnitude of the anomeric effect and consequently O—P bond strengths. Changes in lone pair donor ability (by protonation or coordination with Mg^{2+}) easily affect the magnitude of the anomeric effect and institute significant changes in O—P bond lengths. This new appreciation of the inherent stereoelectronic effects in a range of important phosphoryl compounds provides another factor to be considered in the study of phosphoryl transfer reactions in nature.

STEREOELECTRONIC EFFECTS AND THE REACTIVITY OF PHOSPHATES

Introduction

Chapter 4 articulated the possibility that the anomeric effect was the origin of the high-energy bond. Here the relationship between the anomeric effect and the rate of phosphate monoester hydrolysis is investigated.

The remarkable proficiency of enzymatic phosphoryl transfer can reach orders of 10^{21}.[15] The ways in which enzymes achieve catalysis are well documented,[138] but discussion of stereoelectronic effects as an efficient catalytic strategy is rare.

Before an assessment of whether phosphoryl transfer enzymes can utilize the anomeric effect as a catalytic strategy it is imperative to first derive a basic understanding of how the anomeric effect impacts enzyme-free, uncatalyzed, phosphoryl transfer reactions. In solution, phosphoryl transfer reactions show significant rate enhancements. Changing the nature of the organic group attached to the phosphates is sufficient to achieve rate enhancements of 10^{12}.[15] Changing the nature of the solvent to less polar organic solvents, achieve rate enhancements of 10^5.[139] Understanding the physical origin of such rate enhancements is an important first step to an overall understanding of the catalytic strategies employed by enzymes.

Phosphoryl transfer reactions proceed via three main pathways, dissociative, associative and concerted, Figure 1.3. Dissociative and associative pathways dominate the discussion of phosphate monoester hydrolysis in solution, the subject of this chapter. The concerted pathway involving a phosphorane intermediate has only been found in an enzymatic environment, [18] will therefore be discussed no further in this chapter. Dissociative mechanisms are characterized by full cleavage of the bridging O—P bond and expulsion of a monoanionic metaphosphate intermediate prior to nucleophilic addition, Figure 5.1. Associative mechanisms are distinguished by pentacoordinated transition states that possess almost equal degrees of bond formation with the incoming

nucleophile and bond cleavage with the donor phosphate, Figure 5.1. It should be noted however that this description applies to the extremes of the two mechanisms and that in reality a mix of the two extremes can exist.

ASSOCIATIVE MECHANISM

transition state

DISSOCIATIVE MECHANISM

metaphosphate intermediate
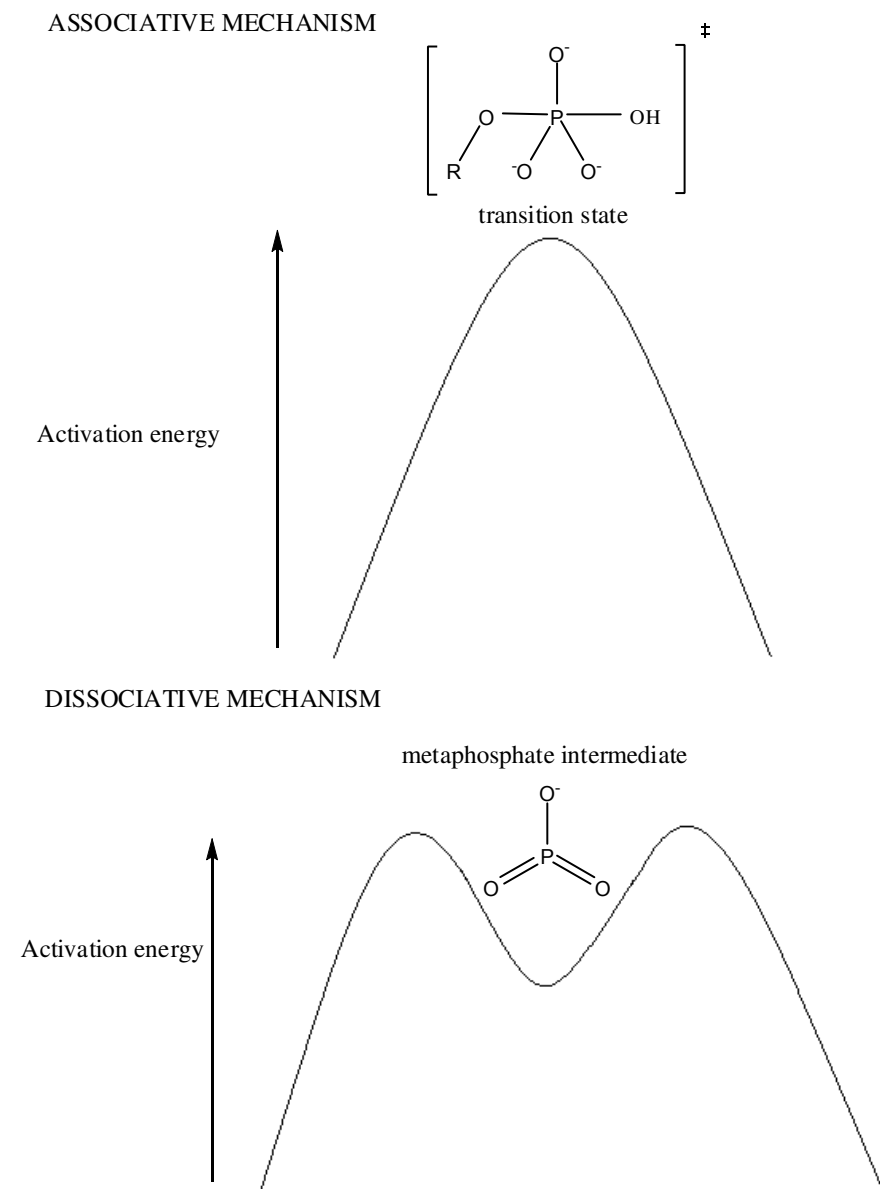
Figure 5.1. Dissociative and associative mechanisms reaction profiles

Phosphate monoester hydrolysis has been the subject of extensive experimental and theoretical work.[15,19,26,34,35,37,38,41,54,70,136,140-152] Perhaps the most characteristic feature of phosphoryl transfer reaction in solution is the correlation between overall hydrolysis rates and the pKa of the leaving group, R

in Figures 1.3 and 5.1. [15,151] This correlation is simply termed a linear free energy relationship or LFER.[15,151] Since pKa's are a measure of electron withdrawing tendencies of the R group, LFER's have long been cited as an indicator of the preference of phosphoryl transfer reactions to adopt the dissociative mechanism in solution.[142] Further experimental information using kinetic isotope effects as a measure of bond dissociation also indicated a preference of solution phosphoryl transfer reactions for a dissociative pathway.[152]

This argument that solution phosphoryl transfer reactions adopt a dissociative pathway has come under severe criticism by theoretical groups.[34,35,38,41] The main criticism was that only determination of an accurate transition structure could provide the insight to determine mechanistic features. Because it was impossible for experimentalists to trap true transition states, it was beyond the capability of existing experiments to provide mechanistic detail. Experiments on ground state structures could not provide enough information about the transition state let alone any strategy employed for lowering its energy. Specifically, experimental methods employed provided a measure of bond dissociation but could not provide any evidence that such bond dissociation led to a decrease in energy for the transition structure. Indeed theoretical modeling associative transition structures in solvent showed that activation energies matched linear free energy profiles and so there was no unique mechanistic interpretation of the linear free energy relationships.[37] However, later theoretical work by Klahn et. al. did show that the mechanism of phosphate hydrolysis in water was more biased towards the dissociative mechanism than the associative mechanism with the associative mechanism only favoured when the pKa of the R group was high.[41] Despite the well-defined reaction pathways and correlation of activation free energies with experimental values described in this work, two important mechanistic questions remain.

(1) Is there an underlying physical origin of LFER's that explain the rate enhancement by substituent effects? It is necessary to demand that any explanation offered should explain the lowering in energy of the transition structure.

(2) Why is there more of a trend towards associative pathways as pKa's increase – is there a simple rule that can explain this?

In this work these questions are examined in the context of the anomeric effect. First archetypal associative and dissociative mechanisms are examined. Then LFER's, are examined in the context of the anomeric effect.

<u>Methods</u>

All electronic structure calculations were carried out with the Gaussian (GO3) program [75,76] using the computational resources at the FSU School for Computational Science and Information Technology (CSIT) and the Center for Computational Sciences at Duquesne University. The energy minimized structures were located using density functional theory (DFT). Specifically, DFT was implemented by using Becke's three-parameter hybrid (exchange) functional [84] with gradient corrections provided by the Lee, Yang, and Parr [85] (B3LYP). The basis set used included both the Pople style 6-311++G(d,p) [98-100] and 6-311++G(3df,2p) [98-100] basis sets. Contributions due to thermal, vibrational, rotational and translational motions, including zero-point energies, were included separately by standard statistical mechanical procedures available in GO3. Frequency analysis has been used to confirm all stationary points as minima or transition structures and provide thermodynamic and zero-point energy corrections. [101]

Where solvent was modelled, the PCM continuum solvation model [127,134], which has been described in detail [153-157], was used. Briefly in the PCM model, the solvent is treated as an infinite continuum of a specified dielectric. The solute or molecule is placed within a cavity defined as a series of interlocking van der Waals spheres centered on the atoms. The interaction between the solute and solvent, or free energy of solvation, is then calculated as a potential term added on to the molecular Hamiltonian. Specific to the PCM model, this potential term includes electrostatic, dispersive and cavitation terms. Because self-consistency between the solute charge distribution and solvent reaction field is required to evaluate the amended Hamiltonian, this method is termed an SCRF or self-consistent reaction field method.

Natural bond order (NBO) [109] analysis was performed using the NBO 3.0 program [42] interfaced into the Gaussian program.

Results and Discussion

The anomeric effect and archetypal associative and dissociative phosphoric transfer reactions. To understand how the anomeric effect impacts phosphoryl transfer reactions archetypal dissociative and associative phosphoryl transfer reactions were first modeled. The associative pathway is distinguished by equal amounts of bond forming with the incoming nucleophile and bond cleavage with the outgoing nucleophile. In the dissociative pathway, full bond cleavage with the outgoing nucleophile occurs prior to bond formation with the incoming nucleophile. As such, the dissociative pathway is characterized with a metaphosphate intermediate unbound to either the incoming or outgoing nucleophile, Figure 5.1 a and 5.1 b. For simplicity the leaving group nucleophiles of both the associative and dissociative reactions were chosen to be methanolate ions. For the associative reaction a methanolate ion was also the donor nucleophile. Because the amount of bond formation and bond cleavage depends on the nature of the nucleophile, the choice of the methanolate ion as both the leaving group and incoming nucleophile assures that the associative transition structure has equal amounts of bond formation and bond cleavage. In both cases, the reaction paths were modeled only to the rate-limiting transition structure. Reactions were modeled in water and acetone, to mimic both solution and enzymatic environments, using the PCM method.

Since at this stage the aim was to examine how the anomeric effect operates in archetypal associative and dissociative phosphoryl transfer reactions, only the dianionic substrate form was considered. The impact of protonation state upon phosphoryl transfer reactions is considered later. The dissociative reaction only involves the expulsion of the monoanionic metaphosphate and the reaction coordinate is simply the distance between the phosphate and the leaving group, Figure 5.2. The reaction coordinate of the associative reaction was more complicated involving both the distance between the phosphate and the leaving

66

group as well as the phosphate and the incoming nucleophile and is also shown in Figure 5.2.

Dissociative
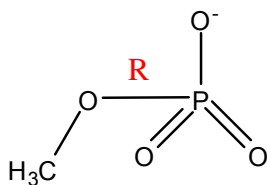
Associative
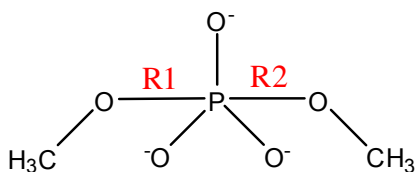
Figure 5.2. Reaction coordinates for associative and dissociative pathways

As described in previous work the NBO method is used to quantify the anomeric effect by treating delocalization from the Lewis structure as a perturbation and employing second-order perturbation techniques to 'measure' the perturbation. Whilst this technique is appropriate for ground-state molecules, transition structures possess a certain caveat that make the use of NBO somewhat inappropriate. Transition structures involve structures whose bonds are breaking and forming. As such there is no clear Lewis structure with which to compare a delocalization in the form of a perturbation to. Thus while the NBO method may be useful in assessing which molecular species is more susceptible to the anomeric effect, the assessment of how the anomeric effect changes along the reaction path may not be best examined using the NBO method. Instead the NRT method is used. The NRT method is based on NBO but instead of assessing delocalization through second-order perturbation methods, provides a percentage weight of resonance structures. Where the anomeric effect is maximum, the % weight of the anomeric resonance structure (herein termed ARS), as shown in Figure 4.6, is also maximized. In this way, the use of a possibly incorrect method of assessing the anomeric effect in transition structures is avoided. It is also

stressed that the anomeric resonance structure as calculated by NRT is somewhat different from resonance structures calculated by other methods. The NRT method calculates the anomeric resonance structure as solely the structure resulting from the $n(O) \rightarrow \sigma^*(O—P)$ delocalization, with loss of the O—P bond but subsequent more double bonded character at the phosphoryl group.

a

Dissociative:Acetone

b

Dissociative:Water

c
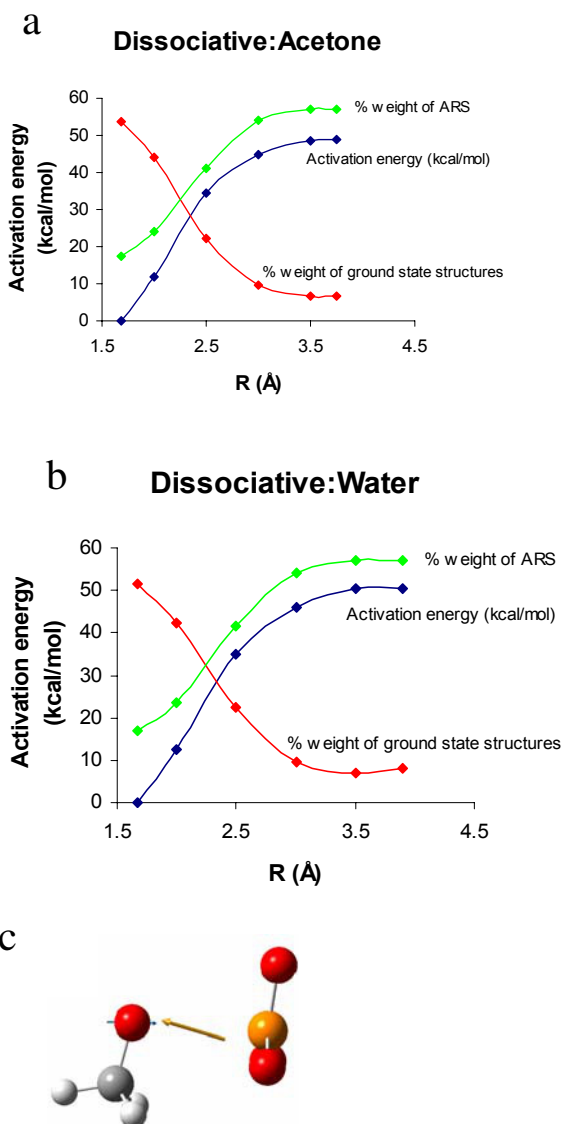
Figure 5.3. Dissociative reaction a) Activation energies, % of anomeric resonance structures and % of ground state structures along the IRC in acetone. b) Activation energies, % of anomeric resonance structures and % of ground state structures along the IRC in water. c) Dissociative transition structure

68

As suspected the % of ARS increases along the reaction coordinate in both water and acetone models and the % ground state structures decrease along the reaction coordinate in both water and acetone models, Figure 5.3. At the transition structure the % of the anomeric, dissociated structure is maximum. This implies that classical dissociative pathways seem to derive their origins from the anomeric effect. Delocalization from lone pairs on the non-bridging oxygen atoms on the phosphate into the antibonding orbital of the bridging O—P bond lengthens that bond. In the dissociative pathway this is essentially cleavage of the O—P bridging bond. A side-effect of the anomeric interaction is that overlap between the lone pairs and σ*(O—P) result in shortened O—P non-bridging bonds. This is basically the formation of the metaphosphate which is characterized by more double-bond character than bonded phosphates.

Because the anomeric effect is a stabilizing effect the maximization of the anomeric effect at the transition structure indicates that dissociative transition states are actually stabilized to some degree by the anomeric effect. Without the anomeric effect imparting a degree of stabilization, dissociative transition structures would be much higher in energy. As often stressed by theoretical groups, the identification of bond weakening by experimental groups alone is not enough to provide a mechanistic determination of phosphoryl transfer. Bond weakening alone does not imply any lowering in energy of the transition structure. The anomeric effect actually provides a mechanism to connect O—P bond weakening with an associated decrease in energy of the transition structure.
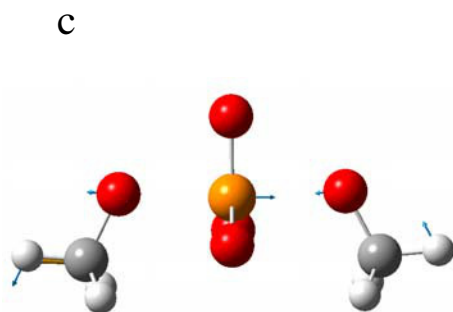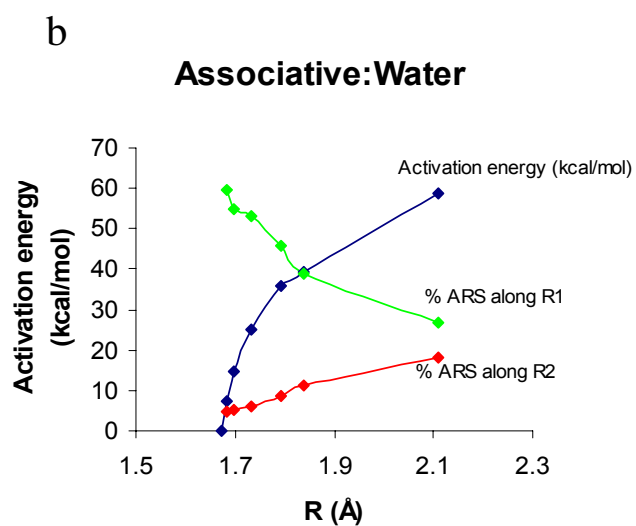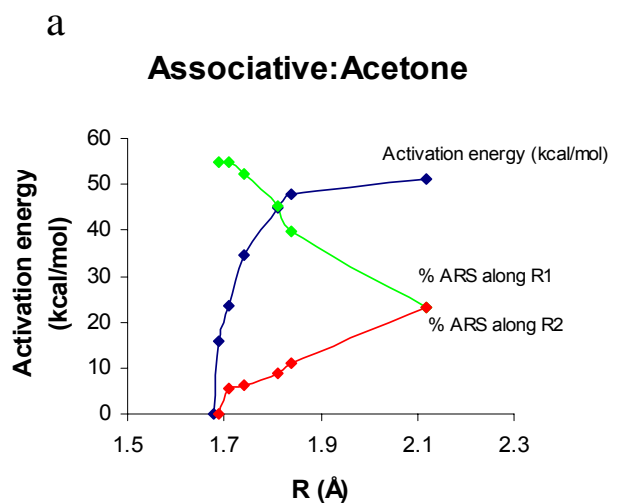
Figure 5.4. Associative reaction a) Activation energies, % of ARS along R1 and R2 in acetone. b) Activation energies, % of ARS along R1 and R2 in water. c) Associative transition structure

In the associative pathway, at the point of the transition structure we find that the % of ARS along coordinate is exactly matched by the % of ARS along the other coordinate, Figure 5.4. Essentially this that the increase in stabilization energy derived from the increase in ARS along one coordinate is matched by the decrease in stabilization energy along the other coordinate. Thus, the associative transition structure derives much less of an energy 'benefit' from the anomeric effect as compared to the dissociative transition structure. As will be shown in later sections, this point is key to the mechanistic preference and the character of the transition structure as the pKa of R increases.

The anomeric effect and linear free energy relationships. 4 selected phosphates with R groups ranging in pKa from 4 to 15.5 were modeled, Table 5.1. The dissociative and associative paths was modeled using various quantum mechanical models, Figures 5.5 and 5.6.

Table 5.1 : Phosphates investigated. Experimental data obtained from Klahn et. al. and references therein[41]

	pK	Expt $\Delta G^{\ddagger}$
methyl monophosphate	15.5	44
phenyl phosphate	10	35
methyl diphosphate	6.3	33
acetyl phosphate	4.8	28
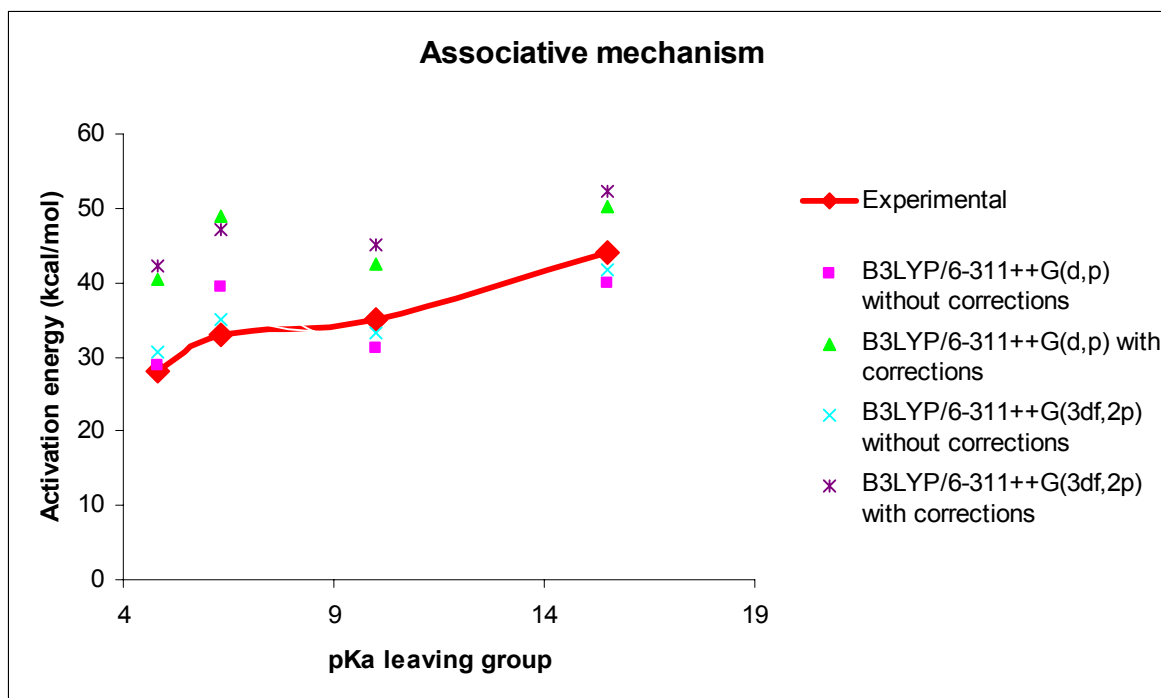
Figure 5.5. Activation energies of the associative reaction.
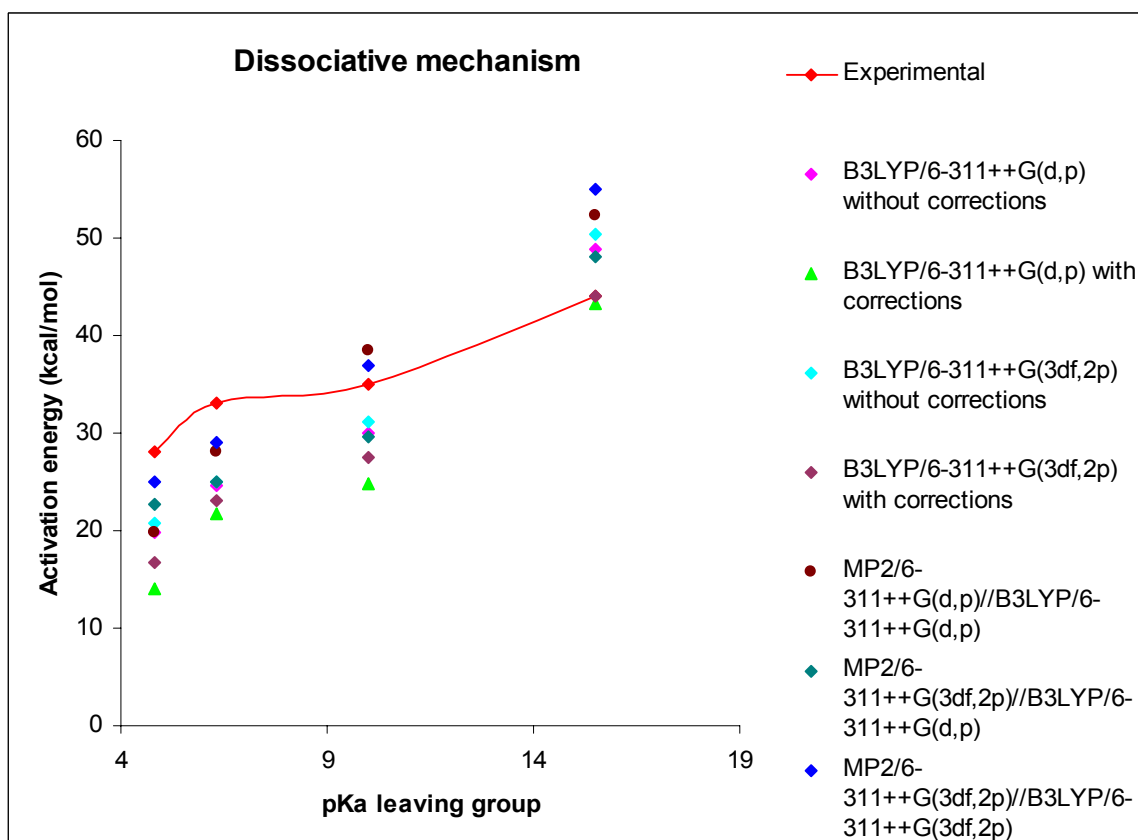
Figure 5.6. Activation energies of the dissociative reaction.

In general, it is found that as discussed before, as the pKa of R decreases, the activation energy of the dissociative mechanism becomes significantly less and so more favourable that the activation energy of the associative mechanism. When the pKa of R increases it becomes harder to distinguish the favoured mechanism as the activation barriers show no clear favourite. We find that calculations of the energy using simple DFT methods together with ZPE corrections do well in reproducing the activation barriers when the pKa of R is high ; at the B3LYP/6-311++G(d,p) level the activation energy of methyl monophosphate is 44.03 kcal/mol almost identical to the experimental value of 44.0 kcal/mol. However, DFT produces significant underestimations when the pKa of R is low. The energy derived from correlated MP2 methods can overcome this but in turn give overestimations of the activation barrier when the pKa of R is high.

The aim of this part is not to derive an accurate description of the reaction pathway since this has already been done so using sophisticated methods by previous groups.[41] Rather, the aim of this section is to investigate the connection between the anomeric effect and rate constants. As described before, an analysis of ground state structures is insufficient to determine the nature of the transition structure and any subsequent energy lowering mechanism thereof. It is also technically incorrect to quantify the magnitude of the anomeric effect on transition structures using E(2) values as the transition structure is made of bonds breaking and forming and so lack any reference Lewis structure. Nevertheless, the main argument of this section is to provide a connection between the anomeric effect and experimental kinetic parameters. The E(2) values of the ground state structures provide a quantification of the magnitude of the anomeric effect in different phosphates and is sufficient to derive this connection. Comparisons between the anomeric effect as quantified in the ground state and experimental activation energies were made as shown in Figure 5.7. As seen, there is a clear correlation between the anomeric effect and experimental activation energies. The increase in the anomeric effect at the

transition structure lowers the energy of the transition structure relative to the ground state and is the underlying origin of rate enhancements seen in LFERs.

Correlation between experimental activation energies and the anomeric effect
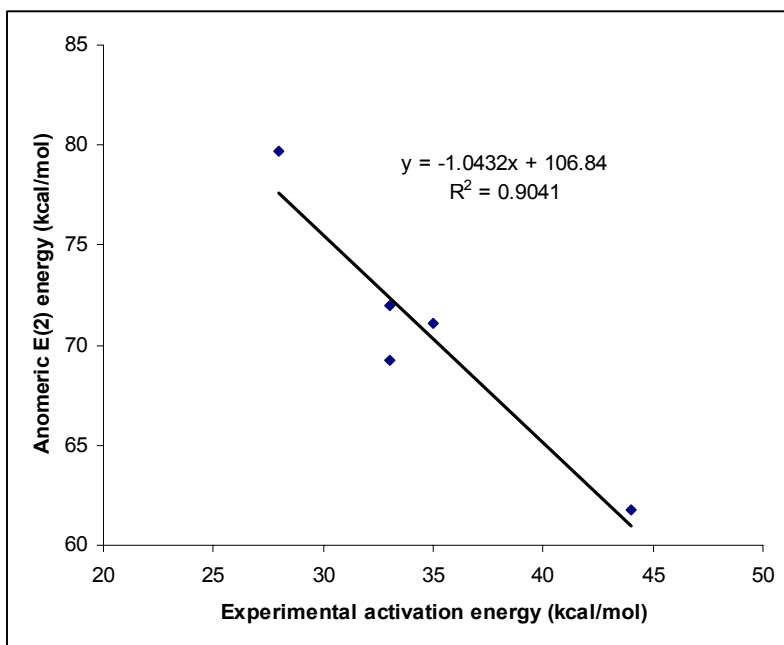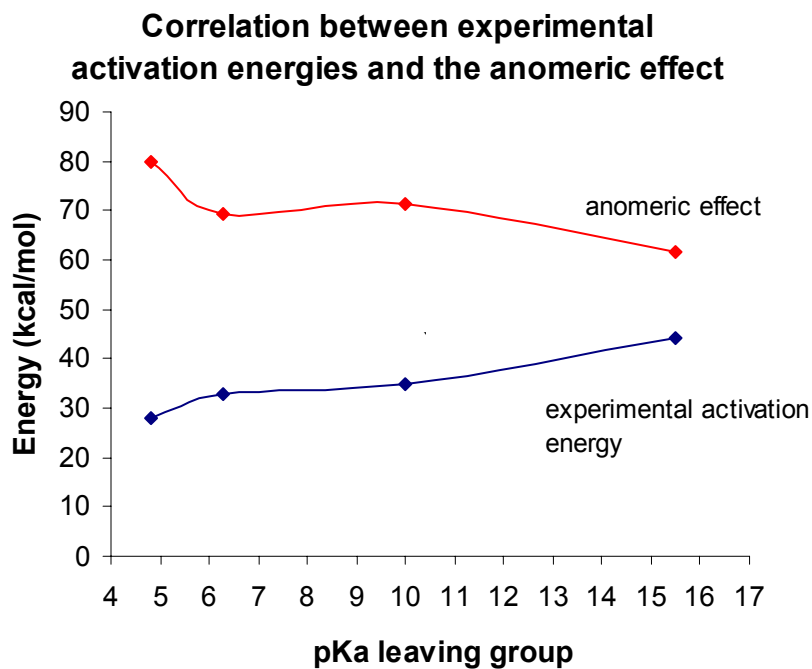
Figure 5.7 Correlation between the anomeric effect and experimental activation energies.

Previous groups have noticed a trend toward the associative pathway as the pKa of R increases.[41] A valence bond approach was used in that work to analyse charge transfer between the ground state and TS. At this point it must be clarified that the VB method of defining charges closely approaches the method of quantifying charge delocalization by NBO. However, the VB method lacks the ability to quantify $n(O)\rightarrow\sigma*(O—P)$ delocalization. Nevertheless, VB charges and sufficient as a necessary check against NBO. It was found that the change in charge transfer in moving from a ground state to a TS did not follow a simple rule. "For example when we have methanolate as a leaving group, the changes are -0.22 and -0.68 for the associative and dissociative TSs, respectively, while for the case of methyl phosphate as a leaving group, the changes are found to be -0.35 and -0.45 respectively. This means that effective TS charges cannot be used to describe the different mechanisms."

The anomeric effect does not rely on isolated charges at the TS rather the drift of charge between the phosphoryl group and the R group. This information, as described by the valence bond method, was provided by previous work, Table 5.3.

Table 5.3. Charges obtained from Klahn et.al.[41]

		R	phosphoryl group	difference
Methyl monophosphate	ground state	-0.29	-1.71	1.42
pKa (R) 15.5	Assoc TS	-0.51	-1.49	0.98
	Dissoc TS	-0.97	-1.03	0.06
Methyl diphosphate	ground state	-1.44	-1.56	0.12
pKa (R) 6.3	Assoc TS	-1.74	-1.26	0.48
	Dissoc TS	-1.89	-1.11	0.78

For the methyl monophosphate ground state, the difference in charge between the R group and the phosphoryl group is high. This indicates little charge drift and corroborates NBO results in that the anomeric effect is reduced when the pKa of R is high. At the dissociative TS, the charges are more equivalent indicating the anomeric effect in operation and charge delocalization between n and $\sigma*(O—P)$ although the VB method lacks the means with which to articulate that. For methyl diphosphate the trend is almost opposite. At the

ground state, charges between R and the phosphoryl group are almost equivalent (-1.44 and -1.56), indicative of the anomeric effect already in action and charge delocalized from the phosphoryl oxygen towards the R group. This is manifested even more at the dissociative TS when the phosphoryl group loses significant charge to the R group leading to a much less negatively charged phosphoryl group (-1.11) compared to the R group (-1.88). For associative mechanisms, a similar yet less extreme trend is derived. There is less charge drift from the phosphoryl groups towards the R group. But R groups with lower pKas display more charge drift toward R the groups with higher pKa.

A simplistic argument is then proposed for the finding that at high R pKa values result in preference for the associative pathway. Because the anomeric effect stabilizes the dissociative transition structure of groups with low R pKa values much more than those with high pKa values, the dissociative transition structure is favoured in groups with low pKa. As this beneficial stabilizing effect is less in groups with high pKas, the associative transition structure becomes more favourable.

Within protein active sites hydrogen bonding, dielectric differences and other electrostatic interactions could also contribute in modulating the magnitude and hence role of the generalized anomeric effect – though perhaps in a less drastic but nonetheless important manner. In arginine kinase, individual residues within the active site were removed and the effect on the O—P bond investigated. It is found that active site residues can enhance or reduce the anomeric effect leading to significant changes in reactive bond lengths, Figure 5.8. It is found that in general arginine residues stabilize the O—P bond and without R229 and R309, the O—P bond is significantly weaker. But this is by no means a general theme with all arginine residues, R126 produces the opposite effect with a stronger O—P bond.

The reason for this can be attributed to the anomeric effect. As seen in figure 5.9, R126 is placed in a position that draws electrons away from the bridging oxygen. This would increase the polarity of the O—P leading to an enhanced anomeric effect and overall weaker O—P bond. Without this residue, the O—P bond is stronger. Interestingly all bond length changes are correlated

with the magnitude of the n(O) → σ*(O—P) interactions. This shows that anomeric effects may play a part in the mechanism and catalysis of phosphoryl transfer reactions and should therefore be investigated.
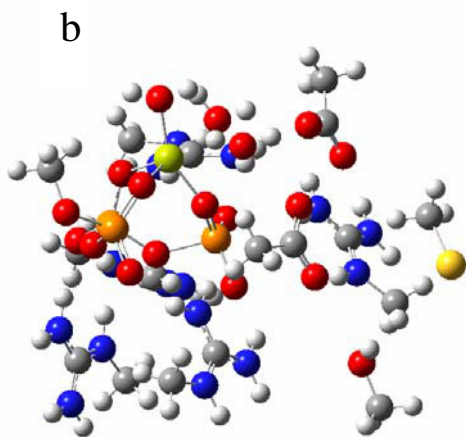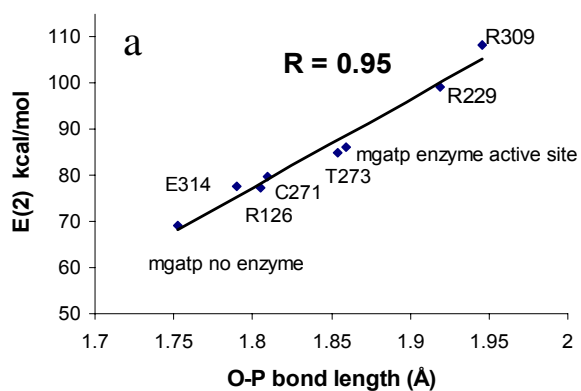
Figure 5.8. a)Correlation between the anomeric effect and O—P bond length in the active site of arginine kinase. Individual residue points refer to optimizations with those residues removed. b) Active site scheme.
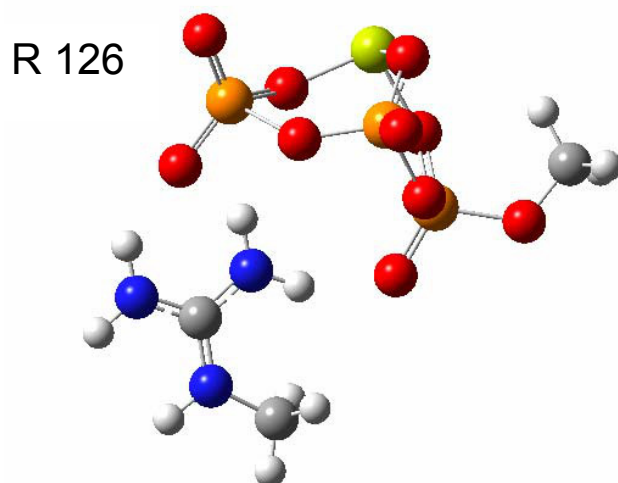
R 126

Figure 5.9. Position of R126

Conclusions

Associative and dissociative phosphoryl transfer mechanisms are described in the context of the anomeric effect. It is found that while the dissociative transition structure benefits from the stabilizing effect of the anomeric interaction, this beneficial effect is diminished in the associative transition structure. In the associative transition structure, the beneficial energy reducing capability of the anomeric effect along the bond dissociative coordinate is offset by a reduction in the anomeric effect along the bond association coordinate. The anomeric effect is found to correlate well with experimental kinetic parameters. The anomeric effect accounts for the finding that the associative pathway is preferred when the pKa of the R group is high. Because the anomeric effect stabilizes the dissociative transition structure of groups with low R pKa values much more than those with high pKa values, the dissociative transition structure is favoured in groups with low pKa. As this beneficial stabilizing effect is less in groups with high pKas, the associative transition structure becomes more favourable.

HYDROGEN BONDING IN PHOSPHATES

Introduction

Metaphosphate, PO_3^-, was first proposed as a key intermediate in the aqueous hydrolysis of phosphate monoesters based upon the independent work by Westheimer and Bunton in 1955.[158,159] Owing to the importance of phosphate hydrolysis to biology, the mechanistic role of PO_3^- in the hydrolysis of phosphate monoesters and other related reactions in solution has since been the subject of intense scrutiny giving way to much experimental and theoretical research.[160-168] Despite extensive studies, several questions remain concerning the nature of PO_3^- compared to other phosphates. Specifically, the difference in stability between PO_3^- and phosphates in water has been well documented by gas phase hydration experiments using mass spectroscopy[160,169,170] and computational investigations,[164,165,171] but a molecular interpretation for this difference in stability remains elusive.

While PO_3^- has been effectively isolated in the gas phase, in clusters with up to three water molecules and in aprotic media, the isolation of PO_3^- in aqueous solution has been difficult.[172] Phosphates on the other hand are stable in solution,[15] which is an important biological property of phosphates, as it allows the compounds to be conveniently utilized for energy storage. Energy release through hydrolysis or phosphoryl transfer occurs only after exposure to an enzyme's catalytic site.[3] The differences in hydrogen bonding between dihydrogen phosphate, $H_2PO_4^-$, and PO_3^- with water molecules may be a key molecular factor in explaining the observed stability differences.

Experimental gas phase hydration studies have shown that that $H_2PO_4^-$ is stabilized by 1.1 kcal/mol over its PO_3^- counterpart in one water complexes

(Scheme 1). This trend in stability follows as more water molecules are added. Disagreement between experiment and computation results for the third hydration step of PO_3^-, where the computed enthalpy of hydration (-7.5 kcal/mol) for the third step is significantly less exothermic than reported by experiment (-16.3 kcal/mol).[164,165,171]
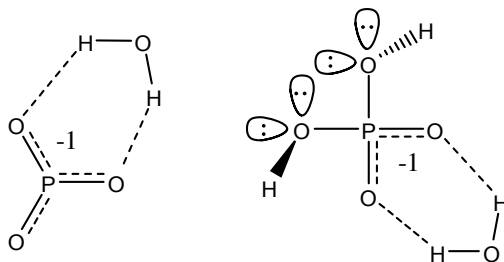
Figure 6.1. Single water complexes of metaphosphate (left, C_{2v}) and dihydrogen phosphate (right, C_1). Hydrogen bonding illustrates the double donor, double acceptor interactions and distal lone pairs of interest.

Water clusters of $H_2PO_4^-$ display more exothermic enthalpies of hydration than PO_3^-. This is unusual, since anionic oxygens are expected to form stronger hydrogen bonds with water.[173-175] $H_2PO_4^-$ has only two unprotonated oxygen atoms ready to accommodate one double hydrogen bond donating water molecule, as shown by Figure 6.1. On the other hand, PO_3^- has three possible locations, so the addition of two and three water molecules should be more favorable than for $H_2PO_4^-$. However, this is not observed by experiment. Thus, water clusters of $H_2PO_4^-$ and PO_3^- are ideal candidates to examine quantum effects that are counterintuitive to simple electrostatic arguments used to explain relative stabilities and hydrogen bonding differences.

Independent theoretical studies using MP2/6-31+G(d,p), CCSD/DZP+ and B3LYP/6-311++G(d,p) levels of theory have been performed for the complexation of PO_3^- and $H_2PO_4^-$ with water molecules.[164,165,171] It was found that low energy complexes formed bifurcated (double donor, double acceptor) hydrogen bonds between PO_3^- and $H_2PO_4^-$ with a single water in symmetric clusters. Computed enthalpies and free energies of hydration from the reports are consistent with experimental values. However, the studies did not uncover the underlying reasons why PO_3^- and phosphates differ in their interaction with water.

Recent experimental data, such as the redshift of infrared X—H stretching modes during hydrogen bond formation,[176] chemical shifts and J couplings in enzyme substrate complexes,[177] and x-ray investigations into hydrogen bonds in ice[178] have suggested that hydrogen bonds involve partial covalent character. The distribution of geometries in protein structures is also indicative of a significant covalent component in many biological environments.[179,180] In recent years, natural bond order (NBO) analysis has been increasingly used to quantify the magnitude of electron delocalization (hyperconjugation) describing the partial covalency of hydrogen bonds. NBO studies have been used successfully to explain the origin of hydrogen bond NMR J couplings in DNA binding,[181] hydroperoxy radicals binding to water surfaces,[113] and the electronic basis of improper hydrogen bonds.[182] NBO analysis has also been used to explain the origin of strong hydrogen bonds.[183] In the NBO methodology, the partial covalency of the hydrogen bond is described by n(lone pair) $\rightarrow$ σ*(antibonding orbitals) hyperconjugation.[184] The magnitude of the n(lone pair) $\rightarrow$ σ*(antibonding orbitals) interaction presents a quantitative assessment of hydrogen bond strength. Motivation to use NBO in this work stems from the need to explain subtle differences in hydrogen bonding for one water complexes, and interpret the unusual asymmetric and strong hydrogen bonds between two or more waters with phosphates. The reported hydration enthalpies cannot be easily explained by electrostatic arguments, and require a detailed investigation of the individual hydrogen bond interactions by NBO.

Due to the physiological importance of phosphates and lack of understanding on how phosphates interact with water, a systematic study of hydrogen bond strength and its relationship to the stereoelectronic structure of PO_3^- and $H_2PO_4^-$ and observed hydration enthalpies has been undertaken. Electronic structure methods with a variety of basis sets, and NBO analysis provide the necessary atomistic and stereoelectronic detail to identify and quantify the origin of differences in hydration enthalpy between PO_3^- and $H_2PO_4^-$ with up to three water molecules. Specifically, n(O) $\rightarrow$ σ*(O—H) hyperconjugation, or partial covalency, is probed to evaluate differences in hydrogen bonding between the phosphates. Central to this work is the

relationship between phosphate conformation, hydrogen bonding configurations, n(O) and σ*(O—H) orbital overlap, charge-transfer magnitude, hydrogen bonding strength, and ultimately the reported hydration enthalpy difference. It is crucial to achieve an understanding of how phosphates interact with water on such an elementary basis before the influence of solvent and pH upon phosphoryl transfer reactions may be better understood.

Methods

All electronic structure calculations were carried out with the Gaussian program.[76] The computational resources were provided by the School of Computational Science and Information Technology (CSIT) at Florida State University and the Center for Computational Sciences (CCS) at Duquesne University.

The electronic description of hydrogen bonding between phosphates and water requires a careful choice of method and basis set.[185] To incorporate the effects of electron correlation, the energy minimized structures were located with density functional (DFT) and second-order Møller-Plesset many-body perturbation (MP2) theories.[83] Specifically, DFT was implemented with Becke's three-parameter hybrid (exchange) functional[84,186] with gradient corrections provided by Lee, Yang, and Parr (B3LYP).[85] Due to the inability of DFT methods to describe dispersive forces,[187-189] MP2 optimizations were carried out on all structures to serve as a point of verification and test of DFT.

The addition of diffuse functions is imperative to describe the spatial distribution of the phosphate anion accurately.[99,190] However, if large numbers of diffuse functions are used, then there is a potential of extra electrons 'escaping'.[191] In addition, an adequate description of polarization is crucial, since it has been shown that p-polarization functions on the hydrogen atoms are important for a variety of hydrogen bonding systems.[192-195] Consequently, twenty basis sets,[95-100,196,197] as shown in Table 6.1, were chosen to evaluate the energetic convergence of the electronic structure studies.

Basis set superposition errors (BSSE) were obtained using the Boys-Bernardi counterpoise correction method.[198] In prior theoretical calculations of

gas phase hydration energies of PO_3^-, BSSE errors were not considered.[164,165,171] It has been argued that BSSE is not necessary, since there is a consistent cancellation between the effects of basis set incompleteness and electron correlation.[164] It has been reported that the inclusion of BSSE shifts computed thermodynamic parameters away from experimental values.[164] In fact, some researchers have concluded that only half of the effects from BSSE should be included.[199] However, it has become increasing clear that the generalization of BSSE importance is difficult and that its impact depends upon the system under investigation and the level of theory implemented.

Contributions due to thermal, vibrational, rotational and translational motions, including zero-point energies, were included separately by standard statistical mechanical procedures available in Gaussian. Frequency analysis was used to characterize all stationary points as minima and provide thermodynamic and zero-point energy corrections at 298 K.[101]

Natural bond order (NBO)[42] analysis was performed using the NBO 3.1 program interfaced into the Gaussian program. NBO transforms the non-orthogonal atomic orbitals from the HF wavefunction into natural atomic orbitals (NAO), natural hybrid orbitals (NHO) and natural bond orbitals (NBO) each of which are complete and orthonormal. This allows electron density to be treated in a more intuitive manner, *i.e.* localized onto bonds and atoms, leading to a better description of the molecule as a localized Lewis structure. In effect, NBO transformation provides filled orbitals that are more concentrated (localized) in terms of occupancies. This then allows delocalizing interactions to be treated as a perturbation through second-order perturbation theory. The E(2) energy values from the second-order perturbation method then provide a reasonable quantitative description of the magnitude of such delocalizing interactions.[22,122]

The NBO method has been cited for overestimating charge-transfer effects[200,201] compared to other methods of decomposing ab initio intermolecular interaction energies. The later includes the Kitaura and Morokuma (KM)[44] and block-localized wavefunction methods.[46,202] For the water dimer, charge-transfer estimates by the KM and NBO methods are -1.8[203] and -9.3[45] kcal/mol. It has been noted that NBO analysis stresses the role of orbital interaction between

filled and unfilled orbitals, whereas the KM analysis emphasizes classical electrostatics from overlapping charge distributions.[47,204] In the application reported here, neither an overestimation by NBO nor underestimation by other methods is critical, because the errors will approximately cancel when examining relative differences. Indeed the similarity of charge-transfer trends between NBO and block-localized wavefunction calculations has been previously noted.[47]

The calculation of interaction enthalpies was done according to the equilibrium reaction used in previous enthalpy of hydration studies for $H_2PO_4^-$ in Equation 18 and PO_3^- in Equation 19 with n = 1 to 3.[164,165]

$$(HO)_2 PO_2 : [H_2O]_{n-1} + H_2O \rightleftharpoons (HO)_2 PO_2 : [H_2O]_n \text{ [eqn 18]}$$

$$PO_3^- : [H_2O]_{n-1} + H_2O \rightleftharpoons PO_3^- : [H_2O]_n \text{ [eqn 19]}$$
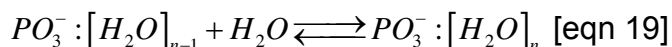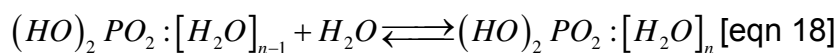
Results and Discussion

The main aim of this work is to understand the difference in the experimentally reported gas-phase hydration enthalpies between PO_3^- and $H_2PO_4^-$ with one to three waters.[164,165,170,205] In particular, the examination of the structural and energetic relationship of the hydrogen bonds formed by different phosphates with water is a crucial first step towards the accurate description of phosphate solvation and reactivity. In this work, a two-step approach has been used to investigate hydrogen bonding in PO_3^- and $H_2PO_4^-$. Firstly, the ability of different levels of theory to reproduce the experimentally determined hydration enthalpies for one water complexes was evaluated. Such studies were used to identify the level of theory needed to converge upon the experimental hydration enthalpy so that similar approaches could be used with larger two and three water systems. Secondly, NBO analysis was utilized to forge a link between subtle structural changes, stereoelectronic effects, and strength of hydrogen bonding at the level of theory producing hydration enthalpies converged to experiment.

Assessment of methods and basis sets with one water phosphate complexes. Multiple starting orientations between the phosphates and a single water resulted in energy minimized bifurcated complexes, as shown in Figure 6.2

As with other theoretical studies, bifurcated complexes were identified as the lowest energy stationary points.[164,165,171] When water was a hydrogen bond acceptor for the hydroxyl group of $H_2PO_4^-$, the hydration enthalpy for the single water complex was far higher than when it was a hydrogen bond donor, so this configuration was not considered further.

mp_1   **hp_1**
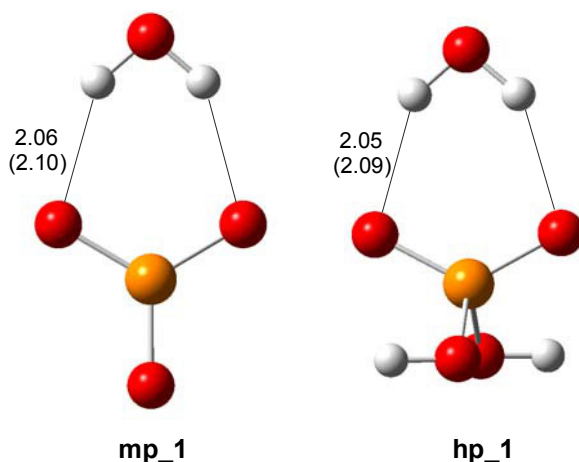
Figure 6.2. Optimized geometries of metaphosphate (**mp_1**), and dihydrogen phosphate (**hp_1**) complexes with one water at the MP2/6-311++G(3df,2p) level. Hydrogen bond lengths are given for MP2/6-311++G(3df,2p) and B3LYP/6-311++G(3df,2p) in parentheses. Bond lengths are in Å. Oxygens are represented by the color red, phosphorous by orange and hydrogen by white.

The complexation of a single water molecule with PO_3^- and $H_2PO_4^-$ were computed at several levels of theory to choose one that can reproduce experimental enthalpies of hydration reported by gas phase high pressure mass spectroscopy[169,170] and electrospray mass spectrometry.[160] Two methods, B3LYP and MP2, in conjunction with up to 16 different Pople style basis sets and four Dunning correlation consistent basis sets were implemented in order to test the incorporation of different functions on heavy and light atoms. The numbers of basis functions of PO_3^- and $H_2PO_4^-$ and interaction enthalpies of the bifurcated complexes at the various levels of theory are given in Table 6.1. Plots of the computed hydration enthalpy as a function of the number of basis functions used in the calculation both with and without BSSE corrections are shown for PO_3^- (Figure 6.3) and $H_2PO_4^-$ (Figure 6.4).

Table 6.1. Number of basis functions of PO_3^- and $H_2PO_4^-$ and interaction enthalpies of the one water complexes at the various levels of theory both with and without BSSE corrections. Experimental enthalpies of interaction are -12.9 +/- 0.3 kcal/mol for PO_3^-;[160] and -14.0 kcal/mol for $H_2PO_4^-$ [169].

			Interaction Enthalpies (kcal/mol)[†]			
Basis Set	**PO_3^-**	**$H_2PO_4^-$**	**PO_3^-**		**$H_2PO_4^-$**	
			no bsse	**bsse**	**no bsse**	**bsse**
B3LYP						
6-31G(d)	83	102	-21.3	-17.5	-18.1	-13.5
cc-pvdz	84	108	-18.5	-11.8	-19.7	-12.3
6-31G(d,p)	89	114	-16.2	-12.5	-20.7	-16.2
6-31+G(d)	103	126	-16.5	-15.4	-14.7	-13.6
6-31+G(d,p)	109	138	-13.5	-12.7	-14.4	-13.6
6-311G(d,p)	110	140	-15.9	-12.0	-17.7	-12.6
6-31++G(d,p)	111	142	-13.8	-13.0	-14.5	-13.6
6-311+G(d)	129	152	-13.6	-12.6	-15	-13.8
6-31G(2d,2p)	125	162	-16.1	-11.4	-17.7	-12.1
6-311++G(d,p)	132	168	-13.3	-12.7	-14.2	-13.4
aug-cc-pvdz	137	178	-12.4	-11.9	-13.1	-12.6
6-311G(2d,2p)	141	182	-15.2	-11.7	-17.1	-12.3
6-31+G(2d,2p)	145	186	-12.5	-12.2	-13.5	-13.0
6-311++G(2d,2p)	163	210	-12.7	-12.1	-13.7	-13.0
cc-pvtz	182	240	-14.3	-12.0	-15.5	-12.8
6-311G(3df,2p)	201	254	-14.5	-11.2	-16	-12.1
6-311++G(3df,2p)	223	282	-12.4	-12.1	-13.4	-13.0
6-311G(3df,3pd)	217	286	-13.6	-11.4	-14.9	-12.1
6-311++G(3df,3pd)	239	314	-12.4	-12.2	-13.5	-13.1
aug-cc-pvtz	280	372	-12.1	-12.1	-13.1	-13.0

Table 6.1

continued

MP2

6-31G(d)	83	102	-16.8	-12.3	-18.4	-12.7
cc-pvdz	84	108	-18.5	-10.9	-19.4	-11.0
6-31G(d,p)	89	114	-16.6	-12.2	-18.1	-12.8
6-31+G(d)	103	126	-14.9	-12.3	-16	-13.1
6-31+G(d,p)	109	138	-14.5	-12.4	-15.5	-13.1
6-311G(d,p)	110	140	-16.3	-11.3	-18	-11.7
6-31++G(d,p)	111	142	-14.6	-12.4	-15.7	-13.0
6-311+G(d)	129	152	-14.6	-12.1	-14.6	-11.6
6-31G(2d,2p)	125	162	-17.2	-11.8	-18.7	-12.4
6-311++G(d,p)	132	168	-14.1	-12.2	-15	-12.8
aug-cc-pvdz	137	178	-13.9	-12.4	-14.9	-13.1
6-311G(2d,2p)	141	182	-16.6	-11.8	-18.5	-12.4
6-31+G(2d,2p)	145	186	-13.8	-12.5	-14.8	-13.4
6-311++G(2d,2p)	163	210	-14.2	-12.5	-15.2	-13.3
cc-pvtz	182	240	-15.6	-12.3	-16.9	-13.3
6-311G(3df,2p)	201	254	-15.9	-11.6	-17.4	-12.7
6-311++G(3df,2p)	223	282	-13.9	-12.7	-15	-13.7
6-311G(3df,3pd)	217	286	-15.4	-12.0	-16.6	-12.8
6311++G(3df,3pd)	239	314	-14.1	-12.8	-15.2	-13.9
aug-cc-pvtz	280	372	-13.9	-13.0	-15.1	-14.1
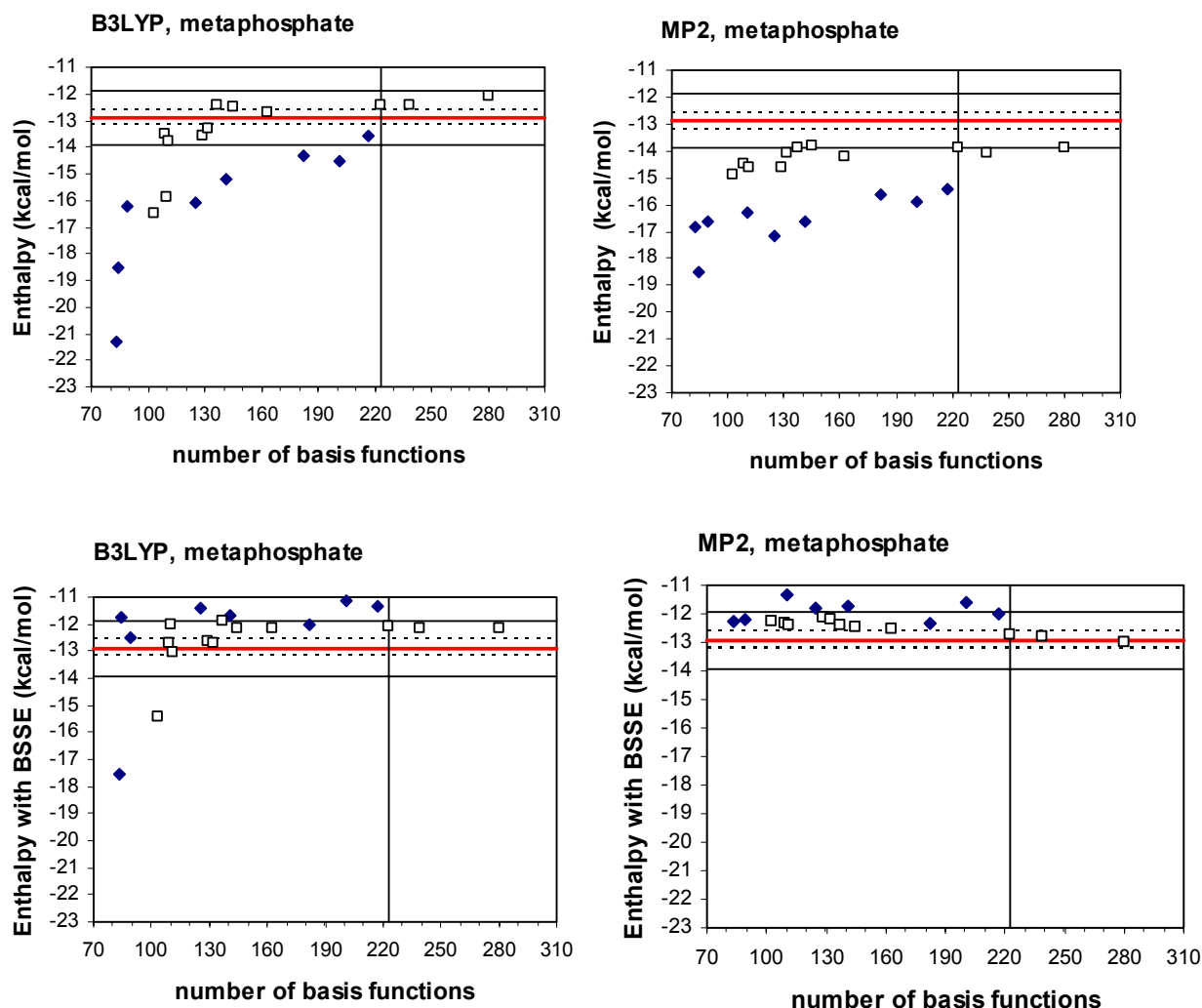
Figure 6.3. A plot of hydration enthalpies versus number of basis functions for the first hydration step of metaphosphate. The experimental enthalpy is given as a solid line in red and calculated values are given in black diamonds (polarization, no diffuse functions) and open squares (polarization and diffuse functions). Vertical lines indicate the 6-311++G(3df,2p) basis set at 223 basis functions selected in this study to represent a converged level of theory. Experimental errors were reported at 0.33 kcal/mol and represented bars as dashed horizontal lines. A reference of 1 kcal/mol is given by solid horizontal lines.

The reported experimental error in the interaction enthalpy between one water and PO_3^- is 0.33 kcal/mol,[160] as indicated by the dashed horizontal lines in Figure 6.3. For visual assistance, a reference of 1 kcal/mol surrounding the experimental value is also given as solid horizontal lines. The experimental error reported for $H_2PO_4^-$ was 1 kcal/mol, as indicated by the dashed horizontal lines in Figure 6.4.[169]

88
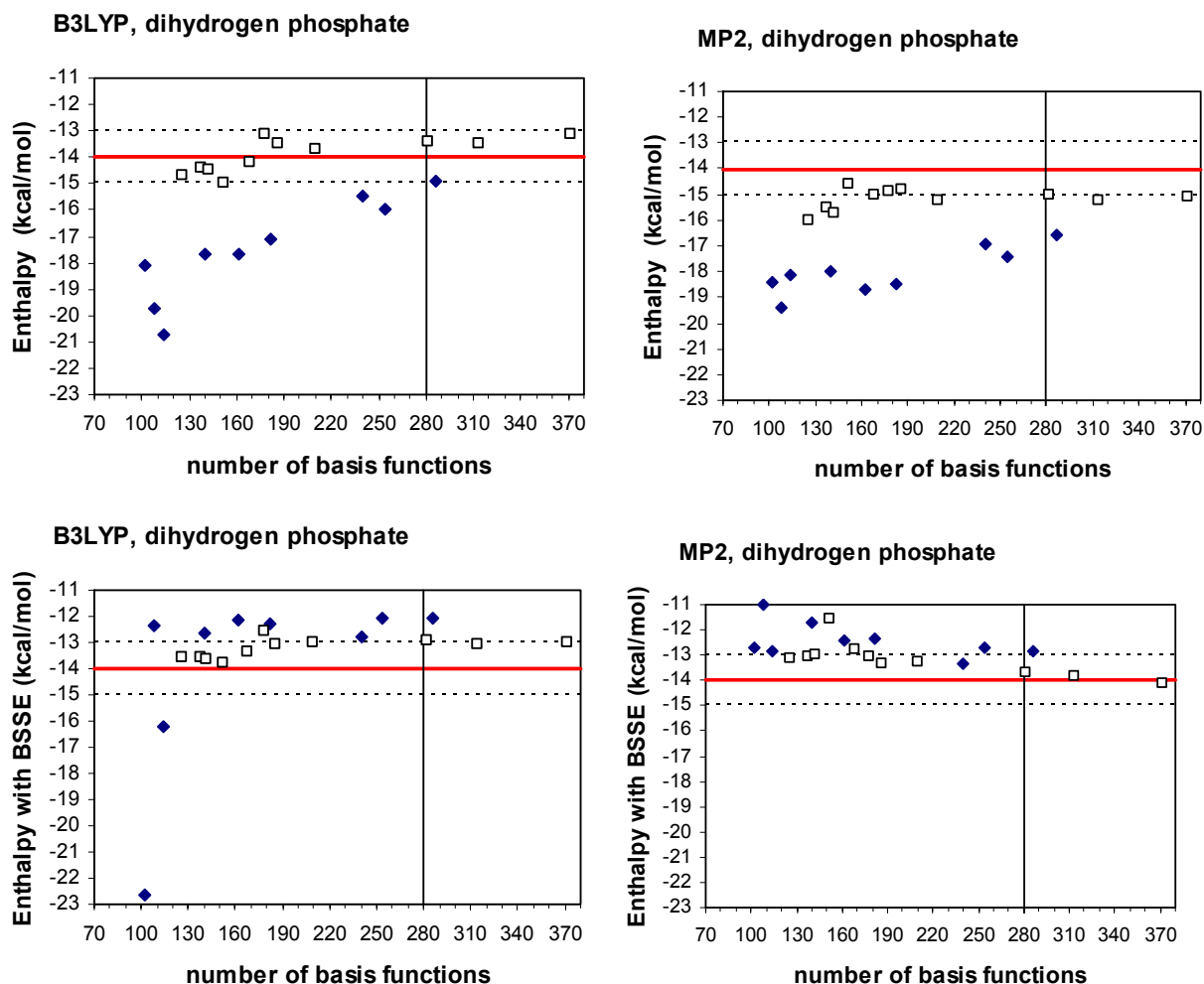
Figure 6.4. A plot of hydration enthalpies versus basis functions for the first hydration step of dihydrogen phosphate. The experimental value is given in red and calculated values are given in black diamonds (polarization, no diffuse functions) and open squares (polarization and diffuse functions). Vertical lines indicate the 6-311++G(3df,2p) basis set at 280 basis functions selected for this study to represent a converged level of theory. Experimental errors were reported at 1 kcal/mol and are represented by dashed horizontal lines.

Previously, the effects of BSSE have largely been ignored in understanding the hydration of phosphate systems.[164,165,171] However, the incorporation of BSSE into the computed hydration enthalpies of these phosphate systems is important to achieve experimental agreement within the reported accuracy, as seen in Figures 6.3 and 6.4. Without BSSE for PO_3^- and $H_2PO_4^-$, both MP2 and B3LYP overestimate the enthalpies of hydration by several kcal/mol. In general, basis sets that rely only upon increases in polarization without BSSE correction (black diamonds, top entries of Figures 6.3 and 6.4) overestimate the enthalpy of hydration. The inclusion of more extensive

polarization functions (without diffuse functions) leads to the improvement of predicted enthalpies of hydration. However, convergence to experimental values is slow using B3LYP, and is never attained using MP2. In essence, the additional polarization does not compensate completely for the incompleteness of the basis set.

In general, basis sets with diffuse functions have a more drastic and rapid impact on the computed enthalpies for both B3LYP and MP2 without BSSE. Basis sets with increases in both diffuse and polarization functions (open squares, top entries of Figures 6.3 and 6.4) produce enthalpies of hydration that are within 1 kcal/mol of experiment when using B3LYP (6-31+G(d) and 6-311G(d,p) are the exceptions for PO_3^-). MP2 without BSSE overestimates the enthalpies by less than 2 kcal/mol consistently across the range of basis sets with diffusion functions. The importance of supplementary functions for the 6-31G(d,p) and 6-311G(d,p) basis sets has been reported[206] for weakly bound complexes[157,207,208] and systems involving lone pair electrons.[160]

Including BSSE in the computations eradicates the large fluctuations in the computed enthalpies of hydration, as seen in Figures 6.3 and 6.4. Improvement in the predicted hydration enthalpies is seen for all levels of theory. The 6-311++G(3df,2p) basis set was selected, since it produced hydration enthalpies within experimental error, and gives virtually identical enthalpies compared to two smaller and two larger basis sets including diffuseness. The basis set chosen delivers experimental accuracy and computational economy of resources. In particular, BSSE brings the computed MP2/6-311++G(3df,2p) into excellent agreement with experiment. The hydration enthalpy of PO_3^- with one water molecule is computed to be -12.7 kcal/mol (-12.9 kcal/mol from experiment),[169] whereas $H_2PO_4^-$ is computed to be -13.9 kcal/mol (-14.0 kcal/mol from experiment).[160,169] The impact of BSSE is less on B3LYP energies than for MP2. However, BSSE has a tendency to over-correct B3LYP energies. The computed B3LYP/6-311++G(3df,2p) hydration enthalpies give slightly better agreement with experiment when basis set superposition is not included. The excellent agreement between experiment and BSSE corrected MP2 results, and strong agreement with B3LYP without BSSE indicates that the orbitals produced

give reliable stereoelectronic insight into the differences of observed complexation enthalpies. Furthermore, it suggests that the complex of two and three water molecules with PO_3^- and $H_2PO_4^-$ will be properly treated using the 6-311++G(3df,2p) basis set with BSSE and MP2. In the treatment of larger systems, the use of B3LYP does not require BSSE.

Orbital Interactions and hydrogen bond strength of PO_3^- and $H_2PO_4^-$ with one water. NBO shows two main hyperconjugative $n(O) \rightarrow \sigma^*(O—H)$ interactions responsible for each hydrogen bond in PO_3^-, but three hyperconjugative $n(O) \rightarrow \sigma^*(O—H)$ interactions responsible for each hydrogen bond in $H_2PO_4^-$, as shown in Table 6.2. The sum of these interactions indicates that hydrogen bonding is stronger in $H_2PO_4^-$, as compared to PO_3^- by 1.3 kcal/mol. To verify the results, deletion energies place this value at 1.2 kcal/mol, a value in good agreement with the experimental difference (1.1 kcal/mol) between enthalpies of PO_3^- and $H_2PO_4^-$ when complexed with one water molecule.[209]

As a percentage of total interaction energies, the E(2) values for $H_2PO_4^-$ and PO_3^- are 76% and 73%. When using deletion energies, the computed percentage reduces to 65% and 61%, respectively. The value of around 60% is in keeping with general NBO evaluations of charge-transfer hydrogen bonding interactions.[209] Nevertheless, both the E(2) and deletion methods give hydrogen bonding, charge-transfer or delocalization energies that are more pronounced in $H_2PO_4^-$ than with PO_3^-.

Energy minimization of metaphosphate and dihydrogen phosphate in the absence of covalent interactions. To test the importance of hyperconjugation effects on the hydrogen bonding differences between $H_2PO_4^-$ and PO_3^-, energy minimizations were performed with the specific $n(O) \rightarrow \sigma^*(O—H)$ interactions removed. In this way, hydrogen bonds formed should be due to the electrostatic contributions and not partial covalency delivered by hyperconjugation. The hydrogen bond lengths are computed to increase from 2.055 Å (with hyperconjugation) to 2.106 Å (without hyperconjugation) for $H_2PO_4^-$, while PO_3^- increases from 2.064 Å (with hyperconjugation) to 2.093 Å (without hyperconjugation), as shown in Table 6.3. This emphasizes the importance of hyperconjugation, because only when its effects are considered do hydrogen

bond lengths and presumably strengths correlate to experiment. Reiterating this point, hyperconjugation provides stronger and shorter hydrogen bonds for $H_2PO_4^-$ compared to PO_3^-.

Table 6.2. Covalent contribution to hydrogen bonding interactions between metaphosphate and dihydrogen phosphate. NBO energy values were calculated using the HF/6-311++G(3df,2p) wavefunction on the MP2/6-311++G(3df,2p) and B3LYP/6-311++G(3df,2p) optimized structures.

		dihydrogen phosphate	metaphosphate
MP2/6-311++G(3df,2p)			
E(2) (kcal/mol)	$n(O) \rightarrow \sigma^*(O\text{-}H)$ 1	3.74	3.89
	$n(O) \rightarrow \sigma^*(O\text{-}H)$ 2	1.01	0.85
	$n(O) \rightarrow \sigma^*(O\text{-}H)$ 3	0.58	
Total E(2) (kcal/mol)		5.33	4.74
Total E(2) of both hydrogen bonds		10.7	9.48
% of total interaction energy		76.1	72.9
Total Edel of both hydrogen bonds		9.06	7.89
% of total interaction energy		64.7	60.6
B3LYP/6-311++G(3df,2p)			
E(2) (kcal/mol)	$n(O) \rightarrow \sigma^*(O\text{-}H)$ 1	2.88	3.43
	$n(O) \rightarrow \sigma^*(O\text{-}H)$ 2	0.85	0.74
	$n(O) \rightarrow \sigma^*(O\text{-}H)$ 3	0.83	
Total E(2) (kcal/mol)		4.56	4.17
Total E(2) of both hydrogen bonds		9.12	8.34
% of total interaction energy		71.0	64.1
Total Edel of both hydrogen bonds		7.74	6.89
% of total interaction energy		60.2	58.3

Table 6.3. Hydrogen bonding in phosphates

	Dihydrogen phosphate	Metaphosphate
Full optimization	2.055	2.064
Optimization without $n(O) \rightarrow \sigma^*(O—H)$ interactions	2.106	2.093

Hydrogen bond orbital interactions between metaphosphate and phosphate. An analysis of the orbitals in both PO_3^- and $H_2PO_4^-$ show that both phosphates share a common motif of interactions, which include primary (defined by the highest E(2) and Edel value) and secondary (defined by the second highest E(2) and Edel value) interactions. The primary $n(O) \rightarrow \sigma^*(O—H)$ interaction, defined by the charge-transfer in both PO_3^- and $H_2PO_4^-$, arises from the interaction between the lone pair $n(O)$ on the phosphate oxygen and $\sigma^*(O—H)$ antibonding orbital of the water directly facing each other, as shown in Figure 6.5.

dihydrogen phosphate   **metaphosphate**
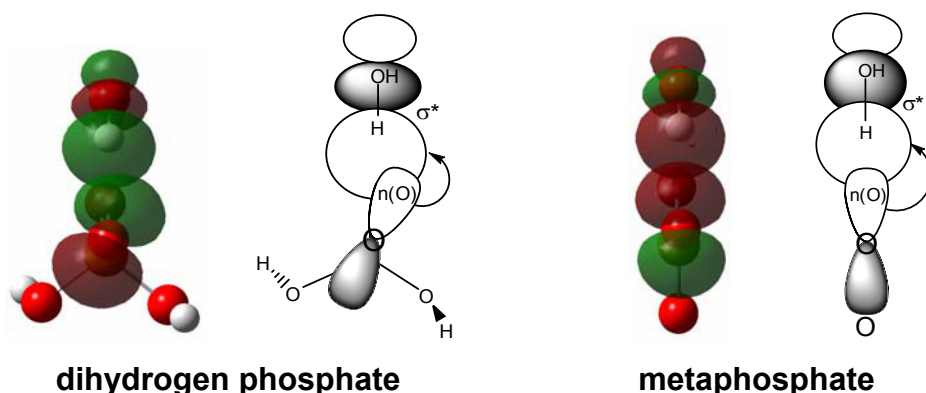
Figure 6.5. Top view of one water complexes with the phosphates illustrating the primary $n(O) \rightarrow \sigma^*(O—H)$ head-to-head interaction between dihydrogen phosphate and metaphosphate with water. For each system, the left-hand figure shows the computed orbitals and the right-hand side is a schematic view.

The two orbitals are oriented head-to-head, providing strong overlap for the primary mechanism for charge-transfer. The secondary $n(O) \rightarrow \sigma^*(O—H)$ interaction in both PO_3^- and $H_2PO_4^-$ arises from the charge-transfer between a

second oxygen lone pair projecting upwards and not directly towards the water σ*(O—H) orbital, as shown in Figure 6.6. The orientation of the two orbitals provides less orbital overlap, as compared to the primary charge-transfer interaction.

dihydrogen phosphate   **metaphosphate**
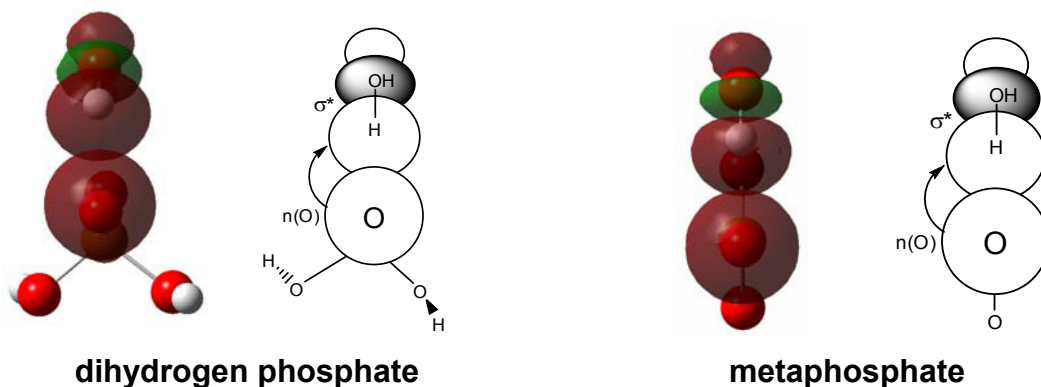
Figure 6.6. Top view of one water complexes with the phosphates illustrating the secondary n(O) → σ*(O—H) interaction. This figure highlights the reduced orbital overlap relative to that shown in the previous figure. For each system, the left-hand figure shows the computed orbitals and the right-hand side is a schematic view.

The final interaction present in $H_2PO_4^-$, but absent in PO_3^-, is crucial to understanding the difference in observed hydration enthalpies, and stems from donation from an oxygen lone pair almost perpendicular to the water's σ*(O—H) orbital. When this lone pair is perfectly perpendicular to the σ*(O—H) orbital, as in the case of PO_3^-, no interaction occurs because both constructive and destructive orbital contributions cancel, as shown in Figure 6.7.

In the case of $H_2PO_4^-$ the donor oxygen lone pair is tilted by ~25°, which is sufficient to allow for orbital overlap and promote charge-transfer, as shown in Figure 6.7. It is computed that the distal lone pair from the hydroxyl substituent clashes with the donor n(O) orbital and causes a distortion that twists the donor lone pair away from the symmetry plane. This is not possible in PO_3^-. The new n(O) orientation creates orbital overlap between the n(O) and σ*(O—H) orbitals resulting in stronger charge-transfer and hydrogen bonding in $H_2PO_4^-$.

dihydrogen phosphate   **metaphosphate**
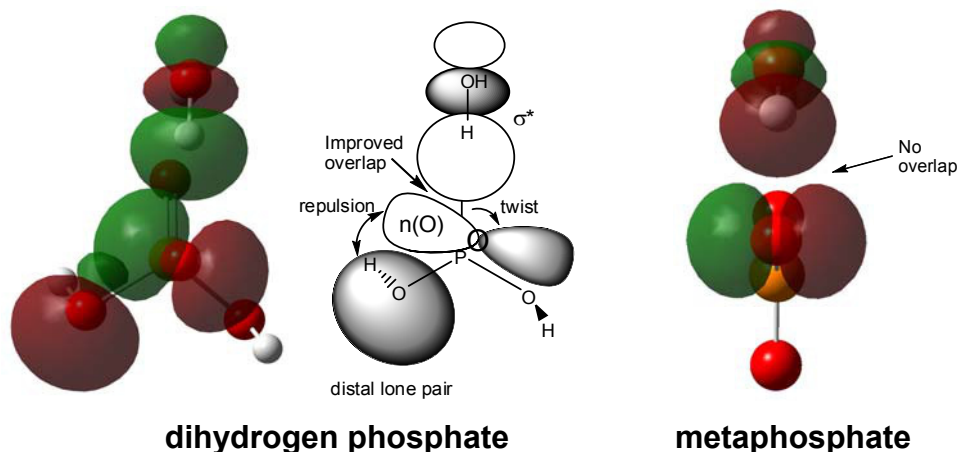
Figure 6.7. Top view of one water complexes with the phosphates illustrating the tertiary n(O) → σ* (O—H) interaction present in dihydrogen phosphate but absent in metaphosphate. For each system, the left-hand figure shows the computed orbitals and the right-hand side is a schematic view.

To further evaluate the hydrogen bond strength dependency on the position of the distal lone pairs, the dihedral of one hydroxyl substituent (H-O-P-O) was rotated in 30 degree increments. As seen in Figure 6.8, the H-O-P-O dihedral dictates the position of the distal lone pairs. As the dihedral is varied, the interaction between the distal lone pair and n(O) changes, which in turn affects the degree of n(O) and σ*(O—H) overlap (co-linearity), and ultimately the hydrogen bond strength.

Starting at a H-O-P-O dihedral of 100° (**A**), a tilt of ~ 25° from co-linearity is observed between the n(O) and σ*(O—H) orbital. By rotating H-O-P-O by 30° from **A** to **B** (Figure 6.8), the clash between the remote lone pair and n(O) becomes more intense, resulting in a 35° tilt in co-linearity between orbitals. The hydrogen bond length increases by 0.01 Å and the charge-transfer reduces by 0.2 kcal/mol. Continued rotation of H-O-P-O diminishes the interaction between the distal lone pair and n(O), resulting with improved co-linearity (**C** through **E**) between n(O) and σ*(O—H) with shortening of the hydrogen bond and strengthening of charge-transfer. Not only does this further validate the previous finding that hydrogen bond strength depends on the position of distal lone pairs, but also that charge-transfer or covalent hydrogen bonding plays a crucial role in hydrogen bond strengths. This finding implies that overall hydration of $H_2PO_4^-$

and PO_3^- depends on distal lone pair orientation, or in simpler terms, hydration depends upon phosphate conformation.
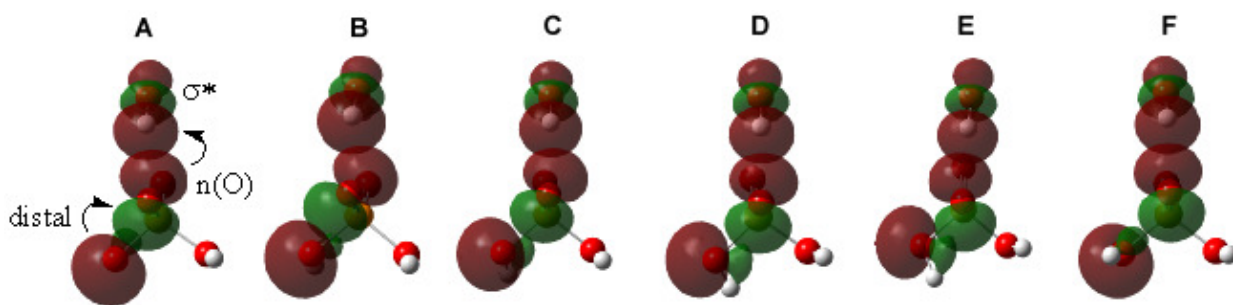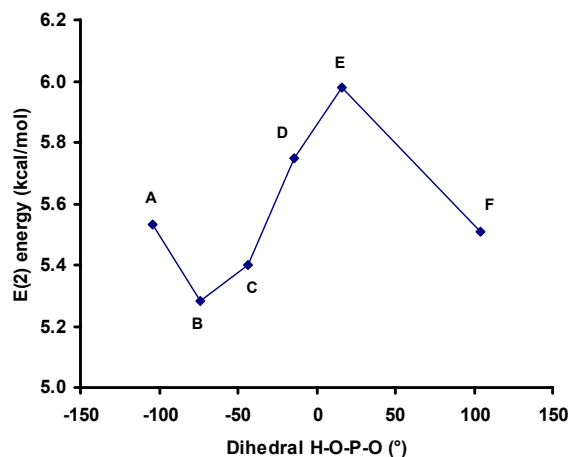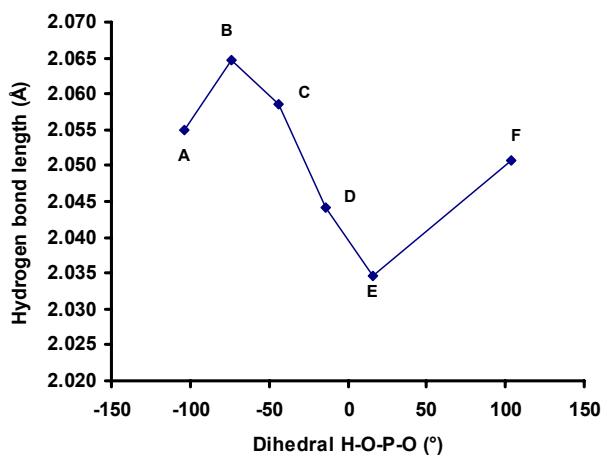
Figure 6.8. Orientation of orbitals as a function of the systematic change of the H-O-P-O dihedral. The impact of distal lone pair clash is shown to affect n(O) → σ*(O—H) co-linearity.

Geometries and energies: Two and three water complexes of metaphosphate and dihydrogen phosphate. Starting orientations for two and

three water complexes with PO_3^- were obtained from prior theoretical calculations.[160,171] Multiple starting orientations were used to search for the low energy complexes of $H_2PO_4^-$ with two and three water molecules. The lowest energy two water configuration is different than that reported earlier[165] being 1.4 kcal/mol lower in energy using B3LYP/6-311++G(3df,2p). The primary difference between the structure previously reported and the newly calculated one is how the two water molecules hydrogen bond with $H_2PO_4^-$. The new structure has one water serving as a double hydrogen bond donor to the deprotonated oxygen atoms, and one water acting as a simultaneous hydrogen bond donor to a deprotonated oxygen and acceptor to a protonated oxygen atom. Both waters in the Houk structure serve as hydrogen bond donors to one protonated and one deprotonated oxygen.[165] The three water structure of $H_2PO_4^-$ has not been previously reported. Despite extensive searching, it cannot be ascertained with certainty if the two or three water complexes with $H_2PO_4^-$ are global minima, although the computed hydration enthalpy for the two water complex is within experimental error. Interaction enthalpies of PO_3^- and $H_2PO_4^-$ complexes with two and three water complexes at the B3LYP/6-311++G(3df,2p) and MP2/6-311++G(3df,2p) level are given in Table 3. Final geometries of the phosphate and water complexes are given in Figure 6.9. Hydrogen bond lengths and the magnitude of $n(O) \rightarrow \sigma^*(O—H)$ interactions are also provided in Figure 6.9.

Hydration enthalpies show that $H_2PO_4^-$ complexes are more stable than their PO_3^- counterparts when complexed with either two or three water molecules.[160,169] The computed MP2/6-311++G(3df,2p) hydration enthalpies corrected for BSSE are within the reported 0.33 kcal/mol experimental uncertainty for PO_3^- and within the 1 kcal/mol uncertainty for $H_2PO_4^-$ for the one and two water complexes. However, the predicted three water PO_3^- complex hydration enthalpy is significantly lower than observed. The large difference between experimental and theoretical hydration enthalpies of PO_3^- with three water molecules has been a long standing source of uncertainty, and addressed elsewhere.[164,165,169]

Table 6.4. Hydration enthalpies of metaphosphate and dihydrogen phosphate at the B3LYP/6-311++G(3df,2p) and MP2/6-311++G(3df,2p) levels of theory corrected for BSSE. Energies not corrected for BSSE are given in parentheses.

| | | Calculated ΔH | | |
	n-1, n	B3LYP	MP2	Expt
Metaphosphate	0,1	-12.1 (-12.4)	-12.7 (-13.9)	-12.9
PO_3^-	1,2	-9.8 (-10.6)	-11.4 (-12.2)	-11.4
	2,3	-7.9 (-9.2)	-7.5 (-10.7)	-16.3
dihydrogen phosphate	0,1	-13.0 (-13.4)	-13.7 (-15.0)	-14.0
$H_2PO_4^-$	1,2	-11.6 (-12.7)	-11.4 (-14.0)	-12.3
	2,3	-12.5 (-14.2)	-8.8 (-13.6)	

To understand the underlying reason for the differences in hydration enthalpies, the hydrogen bonds formed by both PO_3^- and $H_2PO_4^-$ were examined in detail. The main difference between PO_3^- and $H_2PO_4^-$ is that the hydrogen bonds formed between water and $H_2PO_4^-$ are unequal in terms of geometry and strength as compared to PO_3^-. Unlike $H_2PO_4^-$, the hydrogen bonds in the two water complex with PO_3^- are approximately equal in length (2.09 to 2.11 Å). The three water complex has uniform hydrogen bonding distances at 2.13 Å. For all PO_3^- complexes considered, the waters behave as double hydrogen bond donors. In contrast, hydrogen bond lengths in $H_2PO_4^-$ complexes can differ by as much as 0.46 Å. For example, in the two water complex, **hp_2**, hydrogen bond lengths of 2.14, 2.01 and 1.68 Å are found. One water behaves as a double hydrogen bond donor, whereas the other water serves as a hydrogen bond donor and acceptor. In the three water complex, **hp_3**, hydrogen bond lengths of 2.09, 2.02 and 1.70 Å are computed. In this case, one water behaves as a double hydrogen bond donor, whereas the other two waters serve as a simultaneous hydrogen bond donor and acceptor.
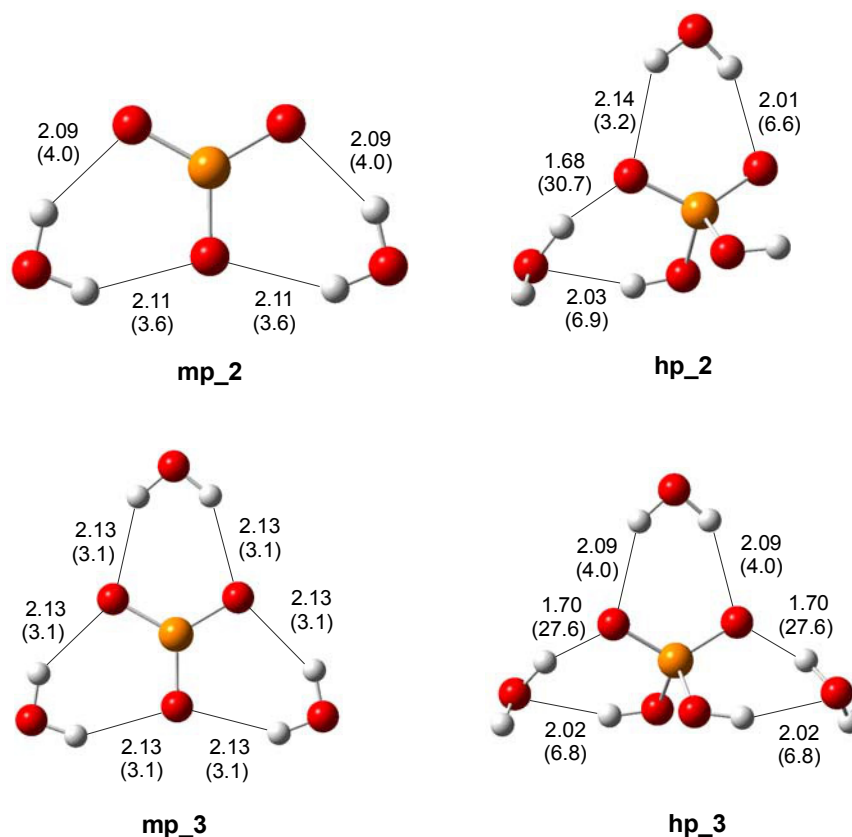
Figure 6.9. Optimized geometries of metaphosphate (**mp_2** and **mp_3**) and dihydrogen phosphate (**hp_2** and **hp_3**) complexes with two and three water at the MP2/6-311++G(3df,2p) with hydrogen bond lengths in Å. E(2) values of the n(O) → σ*(O—H) interactions in kcal/mol are given in parentheses. Oxygens are represented by the color red, phosphorous by orange and hydrogen by white.

Interestingly, an oxygen atom of the phosphate molecules serves in three different capacities with water molecules. It may be a hydrogen bond acceptor for one donor, an acceptor for two simultaneous donors, or a hydrogen bond donor. In **mp_2**, two oxygens are acceptors for a single donor, whereas one oxygen is an acceptor for two hydrogen bond donating waters simultaneously. This difference in hydrogen bonding responsibility explains the small differences computed in hydrogen bond lengths and charge-transfer energies. The oxygen coordinated with two hydrogen bond donors is a slightly weaker acceptor. As shown in Figure 6.10 (right-hand side) for the two water complex with PO_3^-, a single lone pair donates into two separate σ*(O—H) orbitals, diluting the n(O) donor density, lowering the n(O) energy, and increasing the n(O) and σ*(O—H) energy gap. Thus, less charge-transfer (3.6 kcal/mol) and a longer hydrogen bond distance (2.11 Å) results, as compared to the oxygen with a single donor

(4.0 kcal/mol and 2.09 Å) illustrated by the other PO_3^- structure in Figure 6.10 (left-hand side).

water as acceptor   n(O) double acceptor

dihydrogen phosphate

n(O) single acceptor   n(O) double acceptor
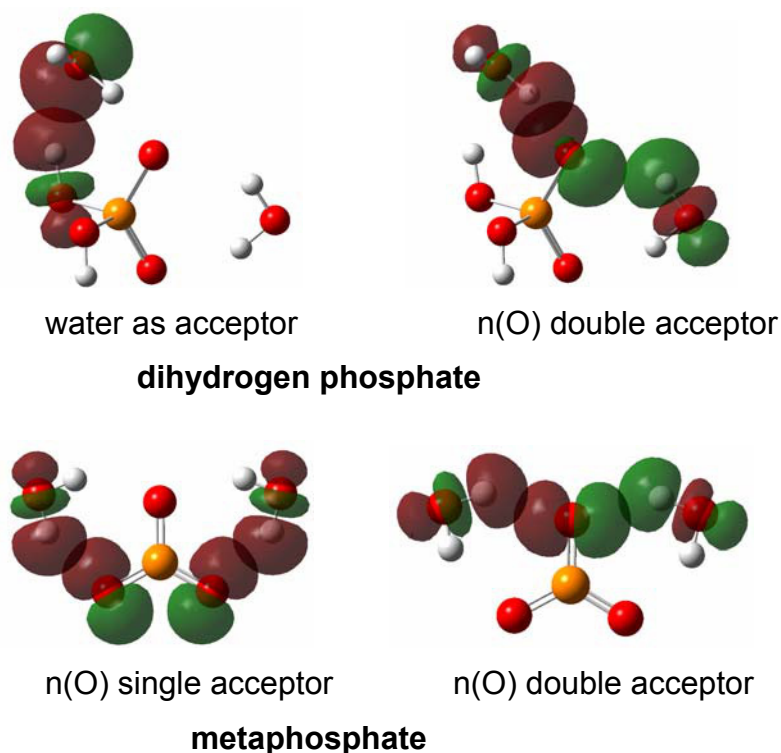
metaphosphate

Figure 6.10.Important orbital interactions of metaphosphate and dihydrogen phosphate with two waters explaining the differences in hydrogen bond strengths. Oxygens are represented by the color red, phosphorous by orange and hydrogen by white.

In **mp_3**, all of the phosphate oxygens are acceptors for two hydrogen bond donating waters simultaneously. Each PO_3^-oxygen atom has a single lone pair donating into two separate σ*(O—H) orbitals (Figure 6.11) resulting in hydrogen bonds weaker than computed in the one and two water complexes for the same reasons stated above. As a consequence of the symmetry, all hydrogen bond lengths (2.13 Å) and charge-transfer energies (3.1 kcal/mol) are equivalent.

Dihydrogen phosphate exhibits different and more asymmetric hydrogen bonding. In the two and three water clusters (Figure 6.9), there is at least one water that serves as a double hydrogen bond donor. In **hp_2**, the same behavior as in the PO_3^- complex is observed, where the phosphoryl oxygen acceptor interacting with a single donor yields stronger hydrogen bonds (6.6 vs. 3.2

kcal/mol) and shorter distances (2.01 vs. 2.14 Å). The three water case **hp_3** produces intermediate distances (2.09 Å) and energies (4.0 kcal/mol) for the formed double donor, double acceptor hydrogen bonds.

| n(O) double acceptor | water as donor | water as acceptor |

dihydrogen phosphate

n(O) double acceptor
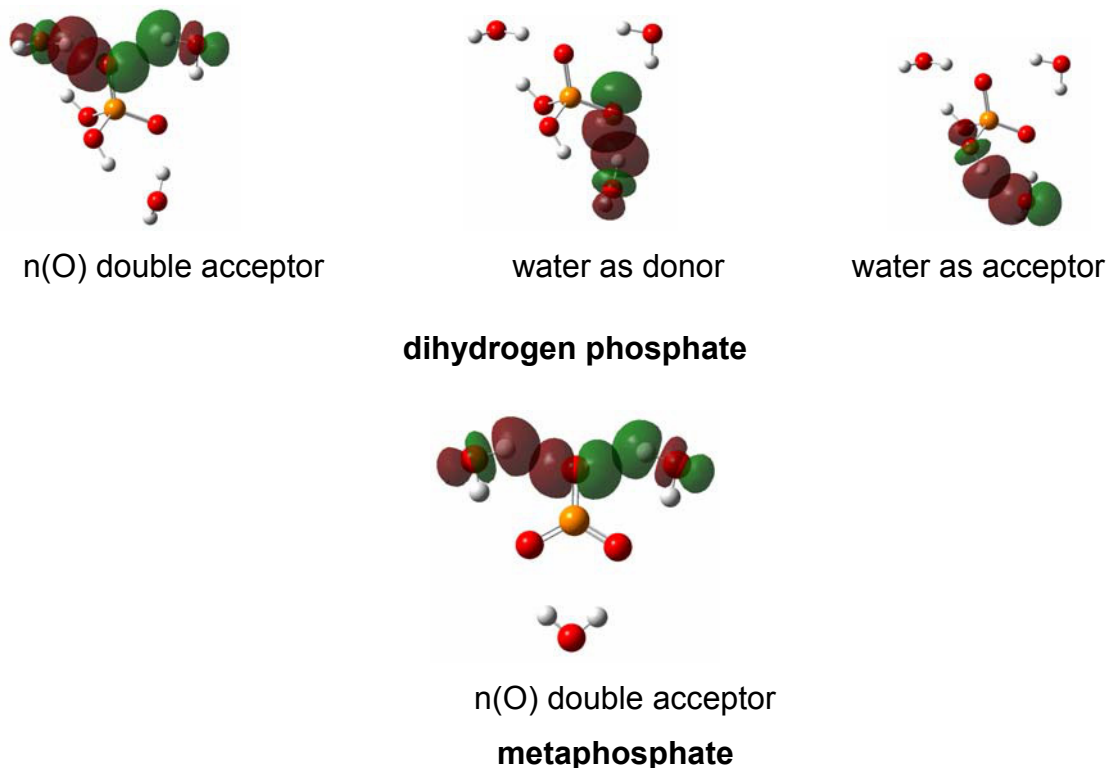
metaphosphate

Figure 6.11. Important orbital interactions of metaphosphate and dihydrogen phosphate with three waters explaining the differences in hydrogen bond strengths. Oxygens are represented by the color red, phosphorous by orange and hydrogen by white.

An interesting complex occurs when a water acts as both a hydrogen bond donor and an acceptor, as found with **hp_2** and **hp_3**. These waters form hydrogen bonds of different strength, even when originating from the same hydrogen bond acceptor atom. When the water is a simultaneous donor and acceptor, one hydrogen bond is extremely short (1.68 and 1.70 Å) between the water donor and oxygen acceptor. Short hydrogen bonds may be key to the observed stability differences between $H_2PO_4^-$ and PO_3^- in two and three water complexes. The NBO method was used to examine the $n(O) \rightarrow \sigma^*(O—H)$ interactions responsible for the short hydrogen bonds. E(2) values corresponding to the respective $n(O) \rightarrow \sigma^*(O—H)$ charge-transfer interactions are provided in Figure 6.9 in parentheses. It is noted that the short hydrogen bonds have

extremely high n(O) → σ*(O—H) values of 30.7 kcal/mol and 27.6 kcal/mol. As previously discussed, these absolute charge-transfer energies are likely over-estimated; however, their relative values indicate that these hydrogen bonds are special and that improved orbital overlap or energies enhance the charge-transfer or partial covalency.

The unusual strength of these short hydrogen bonds stems from the fact that the participating water molecule is also acting as a hydrogen bond acceptor. The NBO orbitals highlighting the different types of donor and acceptor interactions are given in Figures 6.10 and 6.11. Closer examination shows excellent n(O) and σ*(O—H) alignment for cooperative charge-transfer through a single water molecule (Figure 10, center and right-hand side). The withdrawal of electron density from the water molecule reduces the energy of the σ*(O—H) orbital. As such, the computed σ*(O—H) energy is 0.84 a.u., as compared to 0.88 a.u. from the double donor waters on PO_3^- and 0.90 a.u. from the other water molecule on $H_2PO_4^-$. This lowering in energy of the σ*(O—H) reduces the energy gap between n(O) and σ*(O—H) by 25.1 kcal/mol yielding a stronger n(O) → σ*(O—H) charge-transfer interaction and unusually short hydrogen bonds. Overall these findings explain the lack of stability between PO_3^- and water compared to $H_2PO_4^-$.

Biological Implications. Hydrogen bonds are a pervasive interaction in nature and common within enzyme active sites.[17,210-212] How hydrogen bonding, especially short strong hydrogen bonds, impacts catalysis is an important biological question that is actively studied.[140,211-216] The role of this type of hydrogen bonding to enzyme catalysis is particularly interesting, since it has been proposed that such interactions can stabilize transition structures by 8 kcal/mol.[212] In this work for one water complexes, it is shown that distal lone pairs of electrons can perturb orbital orientations and overlap, increase covalency through charge-transfer, and modulate the strengths of hydrogen bonds. In more than one water complexes, charge-transfer can be reduced or greatly magnified depending upon the hydrogen bonding configuration. Since, configurational and conformational change of phosphates are linked to hydrogen bond strength, such changes during active site binding may be crucial to overall enzymatic activity or

phosphoryl transfer reactions. Although $H_2PO_4^-$ is rare in enzymatic systems, other substituted phosphates, rich in lone pairs of electrons, may display the same connection between binding strength and conformational dependency. Future analysis of hydrogen bonding of phosphates in active sites with water, ions or amino acids should include consideration of such stereoelectronic effects.

Conclusions

The electronic basis of hydrogen bonding of metaphosphate and dihydrogen phosphate with one to three waters was studied to understand the experimentally reported differences in enthalpies of hydration. The evaluation of the level of theory utilized shows that it is essential to employ basis set superposition error corrections along with supplemental diffuse functions to compute the enthalpy of hydration to experimental error for both metaphosphate and dihydrogen phosphate. Using natural bond orbital analysis, it is found that dihydrogen phosphate participates in stronger hydrogen bonding relative to metaphosphate due to greater hyperconjugative $n(O) \rightarrow \sigma^*(O—H)$ interactions. The influence of subtle orbital interactions resulting from configurational and conformational changes is key to understanding the hydrogen bonding differences between the two phosphates. In the dihydrogen phosphate complex with one water, the repulsion between distal lone pairs with the $n(O)$ donor orbital creates favorable orbital overlap between $n(O)$ and $\sigma^*(O—H)$, which increases the magnitude of charge-transfer and consequently hydrogen bond strength relative to metaphosphate. In the two and three water complexes, the more asymmetric hydrogen bonding configuration in dihydrogen phosphate allows for cooperative charge-transfer that strengthen its hydrogen bond acceptor ability over PO_3^-. Waters in dihydrogen phosphate complexes serve as both a hydrogen bond donor and acceptor, which reduces the energy gap between $n(O)$ and $\sigma^*(O—H)$, leading to stronger hydrogen bonds. The connection between phosphate conformation, orbital interactions, partial covalency and hydration enthalpies has been demonstrated.

CONCLUSIONS AND FUTURE WORK

Stereoelectronic effects have previously been implicated in phosphates in two ways: (1) Ab initio calculations on small molecule analogs of substrates, that showed the energetic preference for particular stereoelectronic configurations can enhance the reactivity of intermediates. (2) crystal structures of complexes that show stereoelectronic favoured configurations

This dissertation advances this prior knowledge by showing that

1) The anomeric effect is <u>more</u> important than electron withdrawing effects upon the O—P bond in imparting bond weakness.

2) The anomeric effect can account for the thermodynamic poise of a series of phosphates.

3) The anomeric effect is also strongly correlated with rate constants and activation energies for a series of phosphates.

4) Enzymes and other environments can 'tune' the strength of the O—P bond by selectively increasing or decreasing the anomeric effect.

This dissertation has introduced a new appreciation of stereoelectronic effects in phosphates. It has shown that the anomeric effect can explain the thermodynamic instability of phosphates and strongly suggests the role of the anomeric effect in rate enhancements. As a result of these findings scientific discourse on phosphoryl transfer should now include analysis and appreciation of these stereoelectronic factors.

Future work should aim to establish a stronger foundation through direct determination of the change in activation energy and rate attributable to an enzyme's mediation of the stereoelectronic effect. The calculations on the effect of different residues on the O—P bond length initiate this and it should be somewhat trivial to model the reaction with individual residues removed and effects on rate determined. This type of analysis should not be limited to arginine kinase. The investigation of the catalytic effect of stereoelectronics should also not be limited to arginine kinase.

It is not expected that stereoelectronic effects will account for rate enhancements of $10^{12} - 10^{21}$. Rather it is suspected to reach magnitudes to 3-9

kcal/mol or rate enhancements of up to 10^8. The utility of recognizing stereoelectronic effects is that it allows experimentalists to make use of a pre-arranged active site to selectively drive either O—P bond formation or breakage. It can also serve as a tool to allow experimentalists to trap important intermediates within an active site environment.

APPENDIX

American Chemical Society's Policy on Theses and Dissertations
If your university requires a signed copy of this letter see contact information below.

Thank you for your request for permission to include **your** paper(s) or portions of text from **your** paper(s) in your thesis. Permission is now automatically granted; please pay special attention to the implications paragraph below. The Copyright Subcommittee of the Joint Board/Council Committees on Publications approved the following:

> Copyright permission for published and submitted material from theses and dissertations
> ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).
> Publishing implications of electronic publication of theses and dissertation material
> Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journal's editor. If an ACS journal editor considers Web posting to be "prior publication", the paper will not be accepted for publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically.

If your paper has **not** yet been published by ACS, we have no objection to your including the text or portions of the text in your thesis/dissertation in **print and microfilm formats**; please note, however, that electronic distribution or Web posting of the unpublished paper as part of your thesis in electronic formats might jeopardize publication of your paper by ACS. Please print the following credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [JOURNAL NAME], in press (or 'submitted for publication'). Unpublished work copyright [CURRENT YEAR] American Chemical Society." Include appropriate information.

If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in **print or microfilm formats**, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information.

Submission to a Dissertation Distributor: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination through the distributor; ACS's copyright credit line should be printed on the first page of the ACS paper.

Use on an Intranet: The inclusion of your ACS unpublished or published manuscript is permitted in your thesis in print and microfilm formats. If ACS has published your paper you may include the manuscript in your thesis on an intranet that is not publicly available. Your ACS article cannot be posted electronically on a publicly available medium (i.e. one that is not password protected), such as but not limited to, electronic archives, Internet, library server, etc. The only material from your paper that can be posted on a public electronic medium is the article abstract, figures, and tables, and you may link to the article's DOI or post the article's author-directed URL link provided by ACS. This paragraph does not pertain to the dissertation distributor paragraph above.

Questions? Call +1 202/872-4368/4367. Send e-mail to copyright@acs.org or fax to +1 202-776-8112.
10/10/03, 01/15/04, 06/07/06

106

REFERENCES

(1)   Allen, K. N.; Dunaway-Mariano, D. *Trends Biochem Sci* **2004**, *29*, 495.

(2)   Boyer, P. D. *Energy, Life, and ATP - Nobel Lecture*; World Scientific Publishing Co.: Singapore, 1997.

(3)   Voet, D.; Voet, J. G. Biochemistry; 2nd ed.; John Wiley, 1995; pp 428.

(4)   Westheimer, F. H. *Science* **1987**, *235*, 1173.

(5)   Lipmann, F. *Adv Enzymol I* **1941**, *1*, 99.

(6)   Voet, D.; Voet, J. G. *Biochemistry*; John Wiley, 2003.

(7)   Chance, B.; Lees, H.; Postgate, J. R. **1972**, *238*, 330.

(8)   Kammermeier, H. *Basic Res. Cardiol.* **1993**, *88*, 380.

(9)   Garrett, R. H.; Grisham, C. M. Biochemistry; 2nd Edition ed.; Saunders: Fort Worth, 1999; pp 56.

(10)   de Meis, L. *Arch. Biochem. Biophys.* **1993**, *306*.

(11)   Horton, H. R.; Moran, L. A.; Ochs, R. S.; D., R. J.; G., S. K. *Principles of Biochemistry*, 3rd ed.; Prentice-Hall, 2002.

(12)   Kalckar, H. M. *Chem Rev* **1941**, *28*, 71.

(13)   George, P.; Phillips, R. C.; J., S.; Rutman, R. J. *Biochemistry* **1963**, *2*, 508.

(14)   George, P.; Witonsky, R. J.; Trachtman, M.; Wu, C.; Dorwart, W.; Richman, L.; Richman, W.; Shurayh, F.; Lentz, B. *Biochim Biophys Acta* **1970**, *223*, 1.

(15)   Lad, C.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5607.

(16)   *Textbook of Biochemisty with Clinical Correlations*; 5th ed.; Devlin, T. M., Ed.; Wiley-Liss: New York, 2002, pp 1216.

(17)   Zhou, G.; Somasundaram, T.; Blanc, E.; Parthasarathy, G.; Ellington, W. R.; Chapman, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8449.

(18)   Lahiri, S. D.; Zhang, G.; Dunaway-Mariano, D.; Allen, K. N. *Science* **2003**, *299*, 2067.

(19)   Wolfenden, R.; Ridgway, C.; Young, G. *J. Am. Chem. Soc.* **1998**, *120*, 833.

(20)   Abell, K. W. Y.; Kirby, A. J. *Tetrahedron Letters* **1986**, *27*, 1085.

(21)   Grzyska, P. K.; Czyryca, P. G.; Golightly, J.; Small, K.; Larsen, P.; Hoff, R. H.; Hengge, A. C. *Journal of Organic Chemistry* **2002**, *67*, 1214.

(22)   Alabugin, I. V.; Zeidan, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 3175.

(23)   Igor V. Alabugin; Manoharan, M.; Zeidan, T. A. *J. Am. Chem. Soc* **2003**, *125*, 14014.

(24)   Hetenyi, A.; Martinek, T. A.; Lazar, L.; Zalan, Z.; Fulop, F. *J Org Chem* **2003**, *68*, 5705.

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

(26)   Gorenstein, D. G. *Chem. Rev.* **1987**, *87*, 1047.

(27)   Taira, K.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1984**, *106*, 7825.

(28)   Taira, K.; Fanni, T.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1984**, *106*, 1521.

(29)   Florian, J.; Strajbl, M.; Warshel, A. *Journal of the American Chemical Society* **1998**, *120*, 7959.

(30)   Dejaegere, A.; Lim, C.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 4353.

(31)   Cramer, C. J.; Gustafson, S. M. *J. Am. Chem. Soc.* **1993**, *115*, 9315.

(32)   Cramer, C. J.; Gustafson, S. M. *J. Am. Chem. Soc.* **1994**, *116*, 723.

(33)   Banavali, N. K.; MacKerell, A. D. *J Am Chem Soc* **2001**, *123*, 6747.

(34)   Aqvist, J.; Kolmodin, K.; Florian, J.; Warshel, A. *Chemistry and Biology* **1999**, *6*, R71.

(35)   Florian, J.; Aqvist, J.; Warshel, A. *Journal of the American Chemical Society* **1998**, *120*, 11524.

(36)   Glennon, T. M.; Warshel, A. *Journal of the American Chemical Society* **1998**, *120*, 10234.

(37)   Florian, J.; Warshel, A. *Journal of the American Chemical Society* **1997**, *119*, 5473.

(38)   Florian, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719.

(39)   Aqvist, J.; Kolmodin, K.; Florian, J.; Warshel, A. *Chemistry & Biology* **1999**, *6*, R71.

(40)   Holmes, R. R. *Acc. Chem. Res.* **2004**, *37*, 746.

(41)   Klahn, M.; Rosta, E.; Warshel, A. *J. Am. Chem. Soc.* **2006**, *128*, 15310.

(42)   Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO 3.1, 1996.

(43)   Banavali, N. K.; MacKerell, A. D. *J Am Chem Soc* **2001**, *123*, 6747.

(44)   Kitaura, K.; Morokuma, K. *Int. J. Quant. Chem.* **1976**, *10*, 325.

(45)   Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.

(46)   Mo, Y.; Peyerimhoff, S. D. *J. Chem. Phys.* **1998**, *109*, 1687.

(47)   Mo, Y.; Gao, J. *Acc. Chem. Res.* **2007**, *40*, 113.

(48)   Eggleton, P., and Eggleton, G. P. *J. Physiol. (Lond.)* **1927**, *63*.

(49)   James, E.; Morrison, J. F. *Biochim Biophys Acta.* **1966**, *128*, 327.

(50)   Stucki, J. W. *Eur J Biochem* **1980**, *109*, 257.

(51)   Rossi, A. M.; Eppenberger, H. M.; Volpe, P.; Cotrufo, R.; Wallimann, T. *J. Biol. Chem.* **1990**, *265*, 5258.

(52)   Dzeja, P. P.; Terzic, A. *Journal of Experimental Biology* **2003**, *206*, 2039.

(53)   Mildvan, A. S. *Proteins: Structure, Function and Genetics* **1997**, *29*, 401.

(54)   Knowles, J. R. *Annual Review Of Biochemistry* **1980**, *49*, 877.

(55)   Dzeja, P. P.; Bortolon, R.; Perez-Terzic, C.; Holmuhamedov, E. L.; Terzic, A. *PNAS* **2002**, *99*, 10156.

(56)   Lapin, E. P. W., S.; Maker, H.S.; Lehrer, G.M.;Weiss, C. *J. Neurosci. Res.* **1983**, *10*, 9.

(57)   Yousef, M. S.; Fabiola, F.; Gattis, J. L.; Somasundaram, T.; Chapman, M. S. *Acta Crystallogr D Biol Crystallogr* **2002**, *58*, 2009.

(58)   Yousef, M. S.; Clark, S. A.; Pruett, P. K.; Somasundaram, T.; Ellington, W. R.; Chapman, M. S. *Protein Sci* **2003**, *12*, 103.

(59)   Pruett, P. S.; Azzi, A.; Clark, S. A.; Yousef, M.; Gattis, J. L.; Somasundaram, T.; Ellington, W. R.; Chapman, M. S. *J Biol Chem* **2003**, *29*, 26952.

(60)   Gattis, J. L.; Ruben, E.; Fenley, M. O.; Ellington, W. R.; Chapman, M. S. *Biochemistry.* **2004**, *43*, 8680.

(61)   Azzi, A.; Clark, S. A.; Ellington, W. R.; Chapman, M. S. *Protein Science* **2004**, *13*, 575.

(62)   Suzuki, T.; Tomoyuki, T.; Uda, K. *FEBS Lett* **2003**, *533*, 95.

(63)   Schlattner, U.; Forstner, M.; Eder, M.; Stachowiak, O.; Fritz-Wolf, K.; Wallimann, T. *Mol Cell Biochem* **1998**, *184*, 125.

(64)   Schlattner, U.; Eder, M.; Dolder, M.; Khuchua, Z. A.; Strauss, A. W.; Wallimann, T. *Biol Chem* **2000**, *381*, 1063.

(65)   Teague, W. E. J.; Dobson, G. P. *J Biol Chem.* **1992**, *267*, 14084.

(66)   Teague, W. E., Jr.; Dobson, G. P. *J Biol Chem* **1999**, *274*, 22459.

(67)   A *Throughout this text the N--P bond will be referred to as the N'--P bond, consistent with the atom definitions in Figure 1.*

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

(69)   Oesper, P. *Arch Biochem* **1950**, *27*, 255.

(70)   Akola, J.; Jones, R. O. *J. Phys. Chem. B* **2003**, *107*, 11774.

(71)   Wilkie, J.; Gani, D. *J. Chem. Soc., Perkin Trans.* **1996**, *5*, 783.

(72)   Grzyska, P. K.; Czyryca, P. G.; Purcell, J.; Hengge, A. C. *Journal of the American Chemical Society* **2003**, *125*, 13106.

(73)   Hayes, D. M.; Kenyon, G. L.; Kollman, P. A. *J. Am. Chem. Soc* **1978**, *100*, 4331.

(74)   Hill, T. L.; Morales, M. F. *J Am Chem Soc* **1951**, *73*, 1656.

(75)   Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98, revision A.9; Gaussian, Inc.: Pittsburgh, PA,, 1998.

(76)   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; N. Rega; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.;

Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03; Revision D.01 ed.; Gaussian Inc.: Pittsburgh, PA, 2003.

(77)   E *supported by the National Science Foundation, (CHE-0321147) and the Department of Education (P116Z040100)*.

(78)   B *supported by the National Science Foundation (AAB/PSC CHE-030008)*.

(79)   Stewart, J. J. P. *J Comp Chem* **1989**, *10*, 209.

(80)   Stewart, J. J. P. *J Comp Chem* **1989**, *10*, 221.

(81)   Aviyente, V.; Houk, K. N. *J. Phys. Chem. A.* **2001**, *105*, 383.

(82)   Cooney, K. D.; Cundari, T. R.; Hoffman, N. W.; Pittard, K. A.; Temple, M. D.; Zhao, Y. *J. Am. Chem. Soc* **2003**, *125*, 4318.

(83)   Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(84)   Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098.

(85)   Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785.

(86)   Byrd, E. F. C.; Scuseria, G. E.; Chabalowski, C. F. *J. Phys. Chem. B.* **2004**, *108*, 13100.

(87)   Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256.

(88)   Cizek, J. *Adv. Chem Phys* **1969**, *14*, 35.

(89)   Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F. *J. Chem. Phys* **1988**, *89*.

(90)   Scuseria, G. E.; Schaefer, H. F. *J. Chem. Phys.* **1989**, *90*, 3700.

(91)   Gauss, J.; Cramer, C. *Chem. Phys. Lett.* **1988**, *150*, 280.

(92)   Salter, E. A.; W., T. G.; J., B. R. *J. Chem. Phys.* **1989**, *90*.

(93)   Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

(94)   Gordon, M. S.; Binkley, J. S.; Pople J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc* **1982**, *104*, 2797.

(95)   Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797.

(96)   Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(97)   Dill, J. D.; Pople, J. A. *J. Chem. Phys.* **1975**, *62*, 2921.

(98)   Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. *J Chem Phys* **1982**, *77*, 3654.

(99)   Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comp. Chem.* **1983**, *4*, 294.

(100)   Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(101)   Acevedo, O.; Evanseck, J. D. Transition States and Transition Structures. In *Computational Medicinal Chemistry for Drug Discovery*; Bultinck, P., De Winter, H., Langenaeker, W., Tollenaere, J. P., Dekker, M., Eds., 2004; pp 323.

(102)   van Thoai, N. J.; Roche, J. *Ann. N. Y. Acad. Sci.* **1960**, *90*, 923.

(103)   Eggleton, P.; Eggleton, G. P. *Biochem J* **1927**, *21*, 190.

(104)   Ennor, A. H., Morrison, J.F. *Physiol Rev.* **1958**, *38*, 631.

(105) Herriott, J. R.; Love, W. E. *Acta Cryst. B* **1968**, *24*, 1014.

(106) Byrd, E. F. C.; Sherrill, C. D.; Head-Gordon, M. *J. Phys. Chem. A.* **2001**, *105*, 9736.

(107) Pu, J.; Truhlar, D. G. *J. Phys. Chem. A.* **2005**, *109*, 773.

(108) Sheng, Y.; Leszczynski, J. *J. Phys. Chem. A.* **2002**, *106*, 12095.

(109) Reed, A. E.; Weinhold, F.; Curtiss, L. A.; Pochatko, D. J. *J Chem Phys* **1986**, *84*, 5687.

(110) Brunck, T. K.; Weinhold, F. *J. Am. Chem. Soc.* **1979**, *101*, 1700.

(111) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211.

(112) Flores, J. R.; Estevez, C. M.; Carballeira, L.; Juste, I. P. *J. Phys. Chem. A.; (Article); 2001; 105(19); 4716-4725.* **2001**, *105*, 4716.

(113) Belair, S. D.; Hernandez, H.; Francisco, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 3024.

(114) Ulic, S. E.; Oberhammer, H. *J. Phys. Chem. A* **2004**, *108*, 1844.

(115) Cao, X.; Hamers, R. J. *J. Am. Chem. Soc.* **2001**, *123*, 10988.

(116) Tao, F.; Lai, Y. H.; Xu, G. Q. *Langmuir* **2004**, *20*, 366.

(117) Kadossov, E. B.; Rajasekar, P.; Materer, N. F. *J. Phys. Chem. B.* **2004**, *108*, 303.

(118) Kobayashi, J.; Goto, K.; Kawashima, T. *J. Am. Chem. Soc.* **2001**, *123*, 3387.

(119) Kobayashi, J.; Goto, K.; Kawashima, T.; Schmidt, M. W.; Nagase, S. *J. Am. Chem. Soc.* **2002**, *124*, 3703.

(120) Scott, A. P.; Radom, L. *J. Chem. Phys* **1996**, *100*, 16502.

(121) Ruben, E. A.; Chapman, M. S.; Evanseck, J. D. *J Am Chem Soc* **2005**, *127*, 17789.

(122) Alabugin, I. V.; Mariappan, M.; Zeidan, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 14014.

(123) Hwang, M.-J.; Chu, P.-Y.; Chen, J.-C.; Chao, I. *J Comp Chem* **1999**, *20*, 1702.

(124) Meagher, K. L.; Redman, L. T.; Carlson, H. A. *Journal of Computational Chemistry* **2003**, *24*, 1016.

(125) Crowhurst, G. S.; Dalby, A. R.; Isupov, M. N.; Campbell, J. W.; Littlechild, J. A. *Acta Crystallogr D* **1999**, *55*, 1822.

(126) Flachner, B.; Kovari, Z.; Varga, A.; Gugolya, Z.; Vonderviszt, F.; Naray-Szabo, G.; Vas, M. *Biochemistry - US* **2004**, *43*, 3436.

(127) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem Phys* **1981**, *55*, 117.

(128) Singh, U. C.; Kollman, P. A. *J. Comp. Chem.* **1984**, *5*, 129.

(129) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comp. Chem.* **1990**, *11*, 431.

(130) Berg, J.; Tymoczko, J. L.; Stryer, L. Biochemistry; 5th ed.; W.H. Freeman & Co.: New York, 2002; pp 374.

(131) Ma, B.; Meredith, C.; Schaefer, H. F. *J. Phys. Chem.* **1994**, *98*, 8216.

(132) Yang, P.; Murthy, P. P. N.; Brown, R. E. *J. Am. Chem. Soc.* **2005**, *127*, 15848.

(133) Cheng, H.; Nikolic-Hughes, I.; Wang, J. H.; Deng, H.; O'Brien, P. J.; Wu, L.; Zhang, Z.-Y.; Herschlag, D.; Callender, R. *J Am Chem Soc* **2002**, *124*, 11295.

(134) Miertus, S.; Tomasi, J. *Chem Phys* **1982**, *65*, 239.

(135) Stryer, L.; Berg, J. M.; Tmoczko, J. L. *Biochemistry*, 5th ed.; Freeman: New York, 2001.

(136) Barth, A.; Bezlyepkina, N. *J. Biol. Chem.* **2004**, *279*, 51888.

(137) Rutman, R. J.; George, P. *PNAS* **1961**, *47*, 1094.

(138) Fersht, A. *Enzyme Structure and Function*; Freeman: New York, 1985.

(139) Grzyska, P. K.; Czyryca, P. G.; Golightly, J.; Small, K.; Larsen, P.; Hoff, R. H.; Hengge, A. C. *J. Org. Chem.* **2002**, *67*, 1214.

(140) Williams, N. H. *Biochim. Biophys. Acta.* **2004**, *1697*, 279.

(141) Wang, Y. N.; Topol, I. A.; Collins, J. R.; Burt, S. K. *Journal of the American Chemical Society* **2003**, *125*, 13265.

(142) Admiraal, S. J.; Herschlag, D. *Journal of the American Chemical Society* **2000**, *122*, 2145.

(143) Anderson, M. A.; Shim, H.; Raushel, F. M.; Cleland, W. W. *Journal of the American Chemical Society* **2001**, *123*, 9246.

(144) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. *Journal of Physical Chemistry A* **2002**, *106*, 6521.

(145) Benkovic, S. J.; Schray, K. J. Chemical Basis of Biological Phosphoryl Transfer. In *The Enzymes*; Boyer, P. D., Ed.; Academic Press: New York, 1973; Vol. 8; pp 201.

(146) Cavalli, A.; Carloni, P. *J. Am. Chem. Soc.* **2002**, *124*, 3763.

(147) Cleland, W.; Hengge, A. *FASEB J.* **1995**, *9*, 1585.

(148) Dittrich, M.; Hayashi, S.; Schulten, K. *Biophys J* **2003**, *85*, 2253.

(149) Maegley, K. A.; Admiraal, S. J.; Herschlag, D. *PNAS* **1996**, *93*, 8160.

(150) Liu, S.; Lu, Z.; Jia, Y.; Dunaway-Mariano, D.; Herzberg, O. *Biochemistry* **2002**, *41*, 10270.

(151) Kirby, A. J.; Varvoglis, A. G. *J Am Chem Soc* **1966**, *89*, 415.

(152) Grzyska, P. K.; Czyryca, P. G.; Purcell, J.; Hengge, A. C. *J Am Chem Soc* **2003**, *125*, 13106.

(153) Morao, I.; McNamara, J. P.; Hillier, I. H. *J. Am. Chem. Soc.; (Communication); 2003; 125(3); 628-629.* **2003**, *125*, 628.

(154) Cances, E.; Mennucci, B.; Tomasi, J. *The Journal of Chemical Physics* **1997**, *107*, 3032.

(155) Tomasi J.; M., P. *Chem. Rev.* **1994**, *94*, 2027.

(156) Iron, M. A.; Martin, J. M. L.; van der Boom, M. E. *J. Am. Chem. Soc.* **2003**, *125*, 11702.

(157) Freccero, M.; Di Valentin, C.; Sarzi-Amade, M. *J. Am. Chem. Soc.* **2003**, *125*, 3544.

(158) Butcher, W. W.; Westheimer, F. H. *J. Am. Chem. Soc.* **1955**, *77*, 2420.

(159) Bunton, C. A.; Llewellyn, D. R.; Oldham, K. G.; Vernon, C. A. *J. Chem. Soc.* **1958**, 3574.

(160) Keesee, R. G.; Castleman, A. W. *J. Am. Chem. Soc.* **1989**, *111*, 9015.

(161) Henchman, M.; Viggiano, A. A.; Paulson, J. F.; Freedman, A.; Wormhoudt, J. *J. Am. Chem. Soc.* **1984**, *107*, 1453.

(162) Harvan, D. J.; Hass, J. R.; Busch, K. L.; Bursey, M. M.; Ramirez, F.; Meyerson, S. *J. Am. Chem. Soc.* **1979**, *101*, 7409.

(163) Choe, J.-Y.; Iancu, C. V.; Fromm, H. J.; Honzatko, R. B. *J. Biol. Chem.* **2003**, *278*, 16015.

(164) Ma, B.; Xie, Y.; Shen, M.; Schaefer, H. F. I. *J. Am. Chem. Soc.* **1993**, *115*, 1943.

(165) Wu, Y. D.; Houk, K. N. *J. Am. Chem. Soc.* **1993**, *115*, 11997.

(166) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 1654.

(167) Buchwald, S. L.; Friedman, J. M.; Knowles, J., R. *J. Am. Chem. Soc.* **1984**, *106*, 4911.

(168) Calvo, K. C. *J. Am. Chem. Soc.* **1985**, *107*, 3690.

(169) Blades, A. T.; Ho, Y.; Kebarle, P. *J. Am. Chem. Soc.* **1996**, *118*, 196.

(170) Blades, A. T.; Ho, Y.; Kebarle, P. *J. Phys. Chem.* **1996**, *100*, 2443.

(171) Zhang, L.; Xie, D.; Xu, D.; Guo, H. *J. Phys. Chem. A* **2005**, *109*, 11295.

(172) Friedman, J. M.; Knowles, J. R. *J. Am. Chem. Soc.* **1985**, *107*, 6126.

(173) Kropman, M. F.; Bakker, H. J. *Science* **2001**, *291*, 2118.

(174) Kropman, M. F.; Bakker, H. J. *The Journal of Chemical Physics* **2001**, *115*, 8942.

(175) Chandra, A. *J. Phys. Chem. B* **2003**, *107*, 3899.

(176) Thompson, W. H.; Hynes, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 6278.

(177) Arnold, W. D.; Oldfield, E. *J. Am. Chem. Soc.* **2000**, *122*, 12835.

(178) Isaacs, E. D.; Shukla, A.; Platzman, P. M.; Hamann, D. R.; Barbiellini, B.; Tulk, C. A. *Phys. Rev. Lett.* **1999**, *82*, 600 LP

(179) Fabiola, F.; Bertram, R.; Korostelev, A.; Chapman, M. S. *Protein Sci* **2002**, *11*, 1415.

(180) Morozov, A. V.; Kortemme, T.; Tsemekhman, K.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6946.

(181) Wilkens, S. J.; Westler, W. M.; Weinhold, F.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 1190.

(182) Alabugin, I. V.; Manoharan, M.; Peabody, S.; Weinhold, F. *J Am Chem Soc* **2003**, *125*, 5973.

(183) Parreira, R. L. T.; Galembeck, S. E. *J. Am. Chem. Soc.* **2003**, *125*, 15614.

(184) Weinhold, F.; Landis, C. *Valency and Bonding*; Cambridge University Press, 2005.

(185) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389.

(186) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(187) Tuma, C.; Sauer, J. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3955.

(188) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287.

(189) Cybulski, S. M.; Seversen, C. E. *J. Chem. Phys.* **2005**, *122*, 014117.

(190) Modelli, A.; Mussoni, L.; Fabbri, D. *J. Phys. Chem. A* **2006**, *110*, 6482.

(191) Stefanovich, E. V.; Boldyrev, A. I.; Truong, T. N.; Simons, J. *J. Phys. Chem. B* **1998**, *102*, 4205.

(192) Chou, P.-T.; Wei, C.-Y.; Hung, F.-T. *J. Phys. Chem. B* **1997**, *101*, 9119.

(193) Chattopadhyay, S.; Plummer, P. L. M. *Chem. Phys.* **1994**, *182*, 39.

(194) Plummer, P. L. M. *J. Phys. Chem. B* **2004**, *108*, 19582.

(195) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Am. Chem. Soc.* **1992**, *114*, 1645.

(196) Kendall, R. A.; Dunning T. H. Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(197) Woon, D. E.; Dunning, T. H. J. *J. Chem. Phys.* **1993**, *98*, 1358.

(198) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(199) Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. *Chem. Rev.* **2000**, *100*, 4145.

(200) Dreuw, A.; Schweinsberg, H.; Cederbaum, L. S. *J. Phys. Chem. A* **2002**, *106*, 1406.

(201) Torrent-Sucarrat, M.; Sola, M.; Toro-Labbe, A. *J. Phys. Chem. A* **2006**, *110*, 8901.

(202) Mo, Y.; Zhang, Y.; Gao, J. *J. Am. Chem. Soc.* **1999**, *121*, 5737.

(203) Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc.* **1977**, *99*, 1316.

(204) Glendening, E. D.; Streitwieser, A. *J. of Chem. Phys.* **1994**, *100*, 2900.

(205) Blades, A. T.; Klassen, J. S.; Kebarle, P. *J. Am. Chem. Soc.* **1996**, *118*, 12437.

(206) Del Bene, J. E.; Mettee, H. D.; Frisch, M. J.; Luke, B. T.; Pople, J. A. *J. Phys. Chem.* **1983**, *87*, 3279.

(207) Frisch, M. J.; Del Bene, J. E.; Raghavachari, K.; Pople, J. A. *Chem. Phys. Lett.* **1981**, *83*, 240.

(208) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comp. Chem.* **1983**, *4*, 294.

(209) Pendas, A. M.; Blanco, M. A.; Francisco, E. *J. Chem. Phys.* **2006**, *125*, 184112.

(210) Thomaeus, A.; Carlsson, J.; Aqvist, J.; Widersten, M. *Biochemistry* **2007**, *46*, 2466.

(211) Guthrie, J. P. *Chem. Biol.* **1996**, *3*, 163.

(212) Shan, S.; Herschlag, D. *J. Am. Chem. Soc.* **1996**, *118*, 5515.

(213) Perrin, C. L.; Nielson, J. B. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511.

(214) Gerlt, J. A.; Kreevoy, M. M.; Cleland, W. W.; Frey, P. A. *Chem. Biol.* **1997**, *4*, 259.

(215) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem.* **1998**, *273*, 25529.

(216) Czyryca, P. G.; Hengge, A. C. *Biochim. Biophys. Acta.* **2001**, *1547*, 245.

BIOGRAPHICAL SKETCH

Education

BSc (Hons) Biochemistry and Genetics. University of Leeds, U.K. (1997)
Advisor : Dr D. J. Cove
BSc (Hons) Thesis : Isolation and characterization of expressed sequence tags (EST's)
of the moss, Physcomitrella Patens.

MSc Plant Molecular Biology. Southern Illinois University, U.S.A. (2000)
Advisor : Dr D. A. Lightfoot
MSc Thesis : Cloning of the Rhg1 and Rhg4 candidate soybean cyst nematode disease
resistance genes.

PhD Molecular Biophysics. Florida State University, U.S.A. (2007)
Advisor : Dr M. S. Chapman
PhD Thesis : Stereoelectronic effects in phosphates.

Awards

American Heart Association (Florida, Puerto Rico affiliate) pre-doctoral fellowship 2004-
2006 (Ref #:0415212B)

American Heart Association (Florida, Puerto Rico affiliate) pre-doctoral fellowship,
competitive renewal 2006-2007 (Ref #:0615243B)

Kasha award for best graduate student publication, 2006
[*J. Am. Chem. Soc.* 2005 127 (50):17789-17798]

Peer-reviewed Publications

1)Machuka J, Bashiardes S, **Ruben E**, Spooner K, Cuming A, Knight C, Cove D.
Sequence analysis of expressed sequence tags from an ABA-treated cDNA library
identifies stress response genes in the moss Physcomitrella patens.
Plant Cell Physiol. **1999** 40(4):378-387.

2) Meksem K, Zobrist K, **Ruben E**, Hyten D, Quanzhou T, Zhang HB, Lightfoot DA. Two
large-insert soybean genomic libraries constructed in a binary vector: applications in
chromosome walking and genome wide physical mapping.
Theor. Appl. Genet. **2000** 101(5 – 6):747 - 755

3) Meksem K, **Ruben E**, Hyten D, Triwitayakorn K, Lightfoot DA.
Conversion of AFLP bands to high-throughput DNA markers.
Mol. Gen. Genomics. **2001** 265:207-214.

4)Meksem K, **Ruben E**, Hyten D, Schmidt M. and Lightfoot D.A.
High-throughput detection of polymorphism physically linked soybean cyst nematode
resistance gene Rhg4 using Taqman probes.
Mol. Breeding. **2001** 7: 63-71

5) Iqbal J, Afzal J, Yaegashi S, **Ruben E**, Triwitayakorn K, Njiti N, Ahsan R, Wood AJ,
Lightfoot DA
A pyramid of loci for partial resistance to Fusarium solani f. sp. glycines maintains Myo-
inositol-1-phosphate synthase expression in soybean roots.
Theor. Appl. Genet. **2002** 105(8):1115-1123

6) Gattis JL, **Ruben E**, Fenley MO, Ellington WR, Chapman MS.
The active site cysteine of arginine kinase: structural and functional analysis of partially
active mutants.

Biochemistry. **2004** 43(27):8680-8689.

7) **Ruben EA**, Chapman MS, Evanseck JD
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 (50):17789-17798

8) **Ruben E**, Jamai A, Afzal J, Njiti VN, Triwitayakorn, K, Iqbal MJ, Yaegashi S, Bashir, R, Kazi S, Arelli P, Town CD, Ishihara H, Meksem K, Lightfoot DA
Genomic analysis of the rhg1 locus: candidate genes that underlie soybean resistance to the cyst nematode.
Mol. Genet. Genomics. **2006** 276(6):503-516.

Presentations at Scientific Meetings

1) Iqbal MJ, **Ruben E**, Ahsan R, Njiti V, Meksem K, Lightfoot DA
Identification of putative genes conferring resistance to SDS of soybean by differential display of mRNA.
Plant & Animal Genome VIII Conference San Diego, CA, January 9-12, 2000.

2) Meksem K, **Ruben E**, Zobrist KF, Triwitayakorn K, Lightfoot DA
Candidate gene sequences of the cyst nematode resistance loci in soybean.
Plant & Animal Genome VIII Conference San Diego, CA, January 9-12, 2000.

3) Meksem K, **Ruben E**, Zobrist KF, Jamai A, Triwitayakorn K, Arelli PR, Lightfoot DA
Cloning rhg1 in soybean.
Plant & Animal Genome IX Conference San Diego, CA, January 13-17, 2001.

4) **Ruben EA**, Chapman MS, Evanseck, JD

Generalized anomeric effects in *high energy* bonds of phosphagens and triphosphates.

American Chemical Society 231st National Meeting Atlanta, GA, March 26-30, 2006

5) **Ruben EA,** Chapman MS, Evanseck, JD

Stereoelectronic effects in phosphates.

American Chemical Society 233rd National Meeting Chicago, IL, March 25-29, 2007