The Beauty of the Bitter Devils

A Theoretical Study on Phosphate Molecules

Maria Rudbeck
Cover Image: The process of a muscle cell illustrated by Mikael (8 years), Elisa (5), Benjamin (5) and Gabriel (4).

©Maria Rudbeck, Stockholm 2011
Andreas, världens bästa brosa
Abstract

Phosphate transfer reactions are catalyzed by a large number of enzymes comprising kinases, mutases and phosphatases. These enzymes play a fundamental role in controlling numerous life processes and it is therefore important to understand the origin of their potent catalytic power. An example is the Ca$^{2+}$-ATPase. In the E2P-state, this enzyme hydrolyses the phosphorylated amino acid, Asp351, $10^6$ to $10^7$ fold faster than when the model compound, acetyl phosphate, is hydrolyzed in in water.

This thesis explores the catalytic power of Ca$^{2+}$-ATPase using theoretical method based on quantum mechanics. The studies of this protein were made by performing quantum chemical calculations on models of phosphoric monoesters as well as on the explicit reaction pathway of the hydrolysis. The studies show the importance of electrostatic interactions as well as the role of the specific active site residue Glu183, a residue that acts as a base in the catalytic pathway. Furthermore, based on the calculations, the interpretation of the experimental infrared spectrum of the E2P-state of Ca$^{2+}$-ATPase, could be further elucidated as well as modified.

The experimental infrared spectrum of phosphoenol pyruvate in water has also been elucidated through calculations. This molecule is converted into pyruvate in the last step of the glycolytic pathway, a reaction that is catalyzed by pyruvate kinase (PK). These results further enabled the interpretation of the experimental spectrum of the PK's catalytic reaction.

These two processes, the transport of Ca$^{2+}$ into the sarcoplasmatic reticulum against a concentration gradient and the glycolysis, are two important actions of a muscle cell.
List of Publications


IV M. E. Rudbeck, M. R. A. Blomberg and A. Barth, Infrared Spectrum of E2P of Ca\(^{2+}\)-ATPase, *manuscript in preparation*


Reprints were made with permission from the publishers.

Publication not included in this thesis:
Contents

List of Publications.............................................................................................................

Abbreviations......................................................................................................................

1. Introduction....................................................................................................................
   1.1 Glycolysis..............................................................................................................
   1.2 Sarcoplasmic reticulum Ca\(^{2+}\)-ATPase .........................................................

2. Theoretical Background.................................................................................................
   2.1 Quantum Mechanics............................................................................................... 9
      2.1.1 The Hartree-Fock (HF)................................................................................ 11
      2.1.2 Density Functional Theory (DFT)............................................................. 12
      2.1.3 Hybrid Hartree-Fock/Density Functional Method..................................... 14
      2.1.4 Basis Sets....................................................................................................... 16
   2.2 Solvation Effects...................................................................................................... 18
   2.3 Vibrations................................................................................................................ 20
   2.4 Natural Bond Orbital Theory................................................................................ 22

3. Benchmark Studies (Paper II)...................................................................................... 25
   3.1 Introduction.............................................................................................................. 25
   3.2 Basis Set Effects on Wavenumbers...................................................................... 26
   3.3 Solvation Effects on Wavenumbers...................................................................... 28

4. The Infrared Spectra of Phosphoenol Pyruvate (Paper I)......................................... 31
   4.1 Normal Mode Assignment of PEP........................................................................ 32
   4.2 Reactivity of PEP from IR-spectroscopy............................................................ 34

5. The Bitter Devil of E2P (Papers III-V).................................................................. 37
   5.1 Simple environment models.................................................................................. 38
   5.2 Protein environment models................................................................................ 41
   5.3 Structural and Energetic Effects (Paper III)...................................................... 43
   5.4 Charge (Paper III)............................................................................................... 44
   5.5 Vibrations (Paper IV).......................................................................................... 46
   5.6 Reaction Mechanism (Paper V).......................................................................... 49

6. Conclusions and Outlook............................................................................................. 53

7. Acknowledgement....................................................................................................... 57

8. Svensk Populärvetenskaplig Sammanfattning....................................................... 60

9. References.................................................................................................................... 63
"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" but "That's funny...""

Isaac Asimov
1. Introduction

Phosphate groups are composed of the two elements, phosphorus and oxygen. Phosphorus (Φωσφόρος), comes from the Greek word, “light-bearer”, and refers to the Morning star (the planet Venus in its morning appearance). In Latin, the Morning star is called Lucifer, who is generally known as the Devil. The word “oxygen” is derived from the two Greek roots ὀξύς (oxys) and -γενής (-genēs), which translates directly to producing sharpness or bitterness. Together, phosphorus and oxygen bring something essential to life – energy. Without phosphate there would be no adenosine triphosphate (ATP), the “molecular unit of currency” of intracellular energy transfer.

At this very moment your heart is pounding, and soon you will probably stand up to go somewhere. For this to occur, your muscles need to contract, and for contraction, calcium ions and energy in the form of ATP are required. The front cover of this book was illustrated by three of my nephews and my niece in a beautiful and pedagogical fashion: for a muscle contraction to occur, an impulse from a nerve needs to reach the muscle cell. This impulse leads to a release of calcium ions (Ca$^{2+}$) from the sarcoplasmic reticulum (SR), where these ions are stored. The calcium will bind to a protein complex that causes structural change so that actin and myosin can bind to each other. The muscle contraction can now take place, as long as energy in the form of ATP is delivered. This ATP stems from both mitochondria and from the glycolytic pathway.\(^1\) Thank you Mikael, Elisa, Benjamin and Gabriel for introducing two very important players in muscle contraction and setting the stage for this thesis.

A common theme for this thesis is that all molecules studied are phosphorylated. Phosphorylation is one of the fundamental cellular regulatory mechanisms. Phosphate transfer reactions are catalysed by a large number of enzymes comprising kinases, mutases and phosphatases. In the glycolytic pathway, glucose is transformed into pyruvate through a series of reactions, see Figure 1.1. In the last step, phosphoenol pyruvate (PEP) is converted
1. Introduction

into pyruvate and ATP is produced. This reaction is catalysed by pyruvate kinase (PK), a reaction that has been studied using infrared spectroscopy.\(^2,3\) In Paper 1, interesting details of the infrared spectrum of PEP are presented.

Ca\(^{2+}\)-ATPase is an ion pump found in muscle cells, and is involved in the transfer of calcium ions from the cytosol to the sarcoplasmic reticulum against a concentration gradient (Figure 1.1). The active site of the Ca\(^{2+}\)-ATPase is modelled in Papers III-V. By studying this protein in connection with models of phosphoric monoesters, important questions about its catalytic power have been answered.

All work presented here is based on theoretical methods, more specifically on the density functional theory. From quantum mechanical calculations, information about structures, energies, spectroscopic properties and orbitals, can be obtained. An introduction to quantum mechanics (QM) and a number of the properties that can be calculated (and which are discussed in this thesis) are given in Chapter 2. QM is applied to many-body systems using approximate methods, methods that commonly require a basis set. It is important to validate both the method and the basis set being used, since significant errors may occur. This can be accomplished by, for example, comparing calculated results with experimental work. In Paper II, several basis sets were evaluated in order to determine their effect on the vibrational spectrum. This study is highly relevant in this thesis since many of the calculations were compared to experimental infrared-spectroscopic results.

1.1 Glycolysis

Glycolysis is a sequence of reactions where one glucose molecule is oxidized into two pyruvate molecules and two ATP molecules. The important purpose of these metabolic reactions is to produce ATP and pyruvate, where pyruvate is used in both aerobic and anaerobic respiration.\(^1\) In aerobic respiration, the products will enter the citric acid cycle of the mitochondria where considerably more ATP will be produced (~30)\(^4\) Some of this energy is used by muscle cells. However, during for example extreme exercise,
1.1 Glycolysis

Oxygen might be fully consumed, thereby exhausting the respiratory pathway. Oxygen being essential for the function of the citric acid cycle, the muscle cells will turn to fermentation to secure their energy source; a pathway that produces smaller amounts of ATP as well as the by-product lactate.

The conversion from glucose to pyruvate proceeds in ten different steps. The final step is catalysed by the enzyme pyruvate kinase (PK), which produces pyruvate from phosphoenolpyruvate (PEP) (Figure 1.1). In Paper I, the infrared spectrum of PEP was studied both experimentally and computationally.

**Figure 1.1** Different components in a muscle cell. When the cell receives a nerve impulse, Ca\(^{2+}\) is released from the sarcoplasmic reticulum (SR). When the actin-myosin filaments are surrounded by Ca\(^{2+}\) and ATP, the muscle will contract. The ATP is delivered from mitochondria and the glycolytic pathway. The Ca\(^{2+}\)-ATPase will, at the expense of one ATP, transfer two Ca\(^{2+}\) from the cytosol into the lumen of the SR; by reducing the Ca\(^{2+}\)-level around the actin-myosin filament, the muscle will relax.
1.2 Sarcoplasmic reticulum Ca\(^{2+}\)-ATPase

Phosphorylated adenosine triphosphatase (P-ATPases) is the common name for homologous enzymes (in both eukaryotes and prokaryotes) that actively pump cations across cell membranes. The term “P-type” is used when an energy-rich, covalently bonded aspartyl phosphate intermediate is formed. Sarco/endoplasmic (SR/ER) reticulum Ca\(^{2+}\)-ATPases (SERCAs) are P-ATPases that transfer Ca\(^{2+}\) from the cytosol of the cell to the lumen of the SR/ER, at the expense of ATP hydrolysis during muscle relaxation. The SR Ca\(^{2+}\)-ATPases are classified into three different isoforms, SERCA1-3, all coded by distinct genes with distinct functions and expressed in different tissues. SERCA1a is the structurally and functionally best studied P-type ATPase and occurs in muscle cells. Figure 1.2 shows the structure of this enzyme in two different states (E1 and E2) where Ca\(^{2+}\) is either bound or absent. Both states include four different structural domains; three in the cytosol (A, N and P) and one in the membrane, M, which consists of 10 α-helices (M1-M10). The A-domain is the smallest domain and works as an actuator or anchor for the largest domain, N. The N-domain is also the region where the nucleotides bind, while the P-domain contains residue Asp351, to which phosphate is bound in the phosphoenzyme form of Ca\(^{2+}\)-ATPase.

The active transport of Ca\(^{2+}\) into the SR is best described by a reaction cycle (Figure 1.3) where two Ca\(^{2+}\) are transported from the cytoplasm into the SR and two to three protons counter are transported. In the E1 conformation, two Ca\(^{2+}\) ions bind, activating the ATPase to accept ATP as a substrate. As ATP phosphorylates the Asp351 residue, the Ca\(^{2+}\) ions become occluded and the phosphorylated high-energy intermediate, Ca\(_2\)E1P, will be formed. Subsequently, a transition will occur where the Ca\(^{2+}\) ions dissociate from their low-affinity binding-sites into the SR. The two Ca\(^{2+}\) ions are then transported against the concentration gradient in exchange for 2-3 H\(^{+}\). During this transition, the adenosine diphosphate (ADP)-insensitive E2P-intermediate is formed. In its place, dephosphorylation will occur in the presence of water. Finally, the regeneration of the high-affinity binding-sites for Ca\(^{2+}\)
1.2 Sarcoplasmic reticulum Ca$^{2+}$-ATPase completes the cycle, which is reversible in all steps.$^{12}$

![Figure 1.2](image)

**Figure 1.2** Structures of the Ca$^{2+}$-ATPase intermediates: (left) with two Ca$^{2+}$ bound, Ca$_2$E1P (PDB: 1TST), and (right) in the absence of Ca$^{2+}$, E2P (PDB:1WP-G). The colors indicate the different structural domains: the transmembrane domain (M) in blue, phosphorylation domain (P) in yellow, the nucleotide domain (N) in green and the actuator domain (A) in red.

The phosphoenzyme intermediate, Ca$_2$E1P, is ADP-sensitive, while the E2P intermediate reacts with water.$^{13}$ This switch of catalytic specificity is required to ensure that ATP is not wasted before Ca$^{2+}$ is transported. Using infrared difference spectroscopy, the phosphorylation reaction can be observed and the intermediates can be distinguished.$^{14}$ To date, empirical relations and the bond valence model are commonly used to correlate the vibrational wavenumbers to bond lengths and bond strengths.$^{15}$ In Paper IV, it is demonstrated how this empirical method can be improved and an example where it cannot be used. In Paper V, the mechanism of the hydrolysis of E2P is studied.
1. Introduction

Figure 1.3 The active transport of Ca\(^{2+}\) into the SR is best described as a reaction cycle, commonly known as the E1/E2-scheme\(^{13}\). All steps are reversible.
“Science cannot solve the ultimate mystery of nature. And that is because, in the last analysis, we ourselves are a part of the mystery that we are trying to solve.”

Max Planck (1858-1947)
(regarded by many as the founder of quantum theory)

~

"We all agree that your theory is crazy, but is it crazy enough?"

Niels Bohr (1885-1962)
2. Theoretical Background

The Copenhagen interpretation\textsuperscript{a}, devised in the years 1924-27, is the earliest representation of an atom based on quantum mechanics. However, two of its founders, Niels Bohr and Werner Heisenberg, never agreed on the mathematical formalism. Instead, in 1926, Erwin Schrödinger gave an alternative formulation of the theory\textsuperscript{16}, using de Broglie's electron-wave description\textsuperscript{17,18}. Schrödinger's wave function is a pictorial description of the electrons, where they are smeared out, unable to be assigned to a defined location. Later the same year, Max Born suggested that the electrons were not smeared out at all, but rather that the wave function corresponded to a probability distribution\textsuperscript{19,20}. A year later Llewellyn Thomas and Enrico Fermi developed a model to approximate the distribution of the electrons in an atom, as a functional of the electronic probability density\textsuperscript{21}. This model is the predecessor of density functional theory.\textsuperscript{b}

2.1 Quantum Mechanics

In quantum mechanics (QM), properties are expressed by the wave function, $\Psi$, which can be obtained by solving the Schrödinger equation. The Schrödinger equation was proposed by Schrödinger 1926 and its time-independent operator form is shown in Equation 2.1.\textsuperscript{24}

\begin{equation}
H \Psi = E \Psi
\end{equation}

\textsuperscript{a} In the book “The Physical Principles of the Quantum Theory” Heisenberg wrote: “On the whole the book contains nothing that is not to be found in previous publications, particularly in the investigations of Bohr. The purpose of the book seems to me to be fulfilled if it contributes somewhat to the diffusion of that 'Kopenhagen Geist der Quantentheorie' [i.e., Copenhagen spirit of quantum theory] if I may so express myself, which has directed the entire development of modern atomic physics.”

\textsuperscript{b} Classical text books are available for detailed reviews on theoretical chemistry in general\textsuperscript{22,23}
2. Theoretical Background

The eigenvalues, $E$, are the energies of the different energy states, described by $\Psi$. The Hamiltonian operator, $H$, is an energy operator corresponding to both the kinetic ($T$) and the potential ($V$) energy.

Due to inter-electronic correlation, the Schrödinger equation cannot be exactly solved for many-body systems. A first step in solving the equation for many-body systems, is the Born-Oppenheimer (BO) approximation, which states that nuclei can be considered stationary compared to the electrons, due to their large relative mass, which results in low, relative velocities. In other words, if any nucleus is moved, the surrounding electrons adjust to the new nuclear configuration instantaneously. This approximation allows the wave function to be written as a product of two wave functions, one for the electrons and one for the nuclei. Since the nuclei are considered stationary, the kinetic energy of the nuclei is neglected and the resulting Hamiltonian is expressed as the electronic Hamiltonian, $H_e$, see Equation 2.2.

$$H_e = T_e + V_{ee} + V_{ne} + V_{nn}$$  \hspace{1cm} (2.2)

$T_e$ corresponds to the kinetic energy of the electrons, $V_{ee}$ to the repulsive potential energy between electrons, $V_{ne}$ to the attractive potential energy between the nuclei and the electrons, and $V_{nn}$ to the repulsive potential energy between the nuclei. $V_{nn}$ is constant within the BO-approximation.

To date, many methods are available for approximate solving of the electronic Schrödinger equation. Two of the most fundamental are the Hartree-Fock Theory (HF) and Density Functional Theory (DFT). HF uses the variation theorem that implies that the energy corresponding to an arbitrary wave function, is always an upper limit to the true energy. Consequently the best description of a wave function for a given configuration will have the lowest energy. The DFT and Hybrid-DFT methods are both based on the fact that the ground-state energy and other properties of a system can be obtained solely from the electron density, $\rho$. 
2.1 Quantum Mechanics

2.1.1 The Hartree-Fock (HF)

The Hartree-Fock (HF) method uses a trial wave function constructed as a single Slater determinant. A Slater determinant, Equation 2.3, consists of single electron wave functions, known as spin orbitals ($\chi$), which are products of a spatial function (molecular orbital) and a spin function ($\alpha$ or $\beta$). The Slater determinant satisfies both the wave function's anti-symmetry (it changes sign upon interchange of any two electrons) and the Pauli exclusion principle (the determinant equals zero if two electrons with the same spin fill the same orbital).

$$\Psi(1, 2, ..., N) = \frac{1}{\sqrt{N!}} \left| \begin{array}{cccc} \chi_1(1) & \chi_2(1) & \cdots & \chi_n(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(N) & \chi_2(N) & \cdots & \chi_N(N) \end{array} \right|$$ \hspace{1cm} (2.3)

The solution to the HF equations (Equation 2.4) is determined by minimizing the energy (the variational theorem) with respect to the choice of spin orbitals.

$$f_i \chi_i = \epsilon_i \chi_i$$ \hspace{1cm} (2.4)

$f_i$ is an effective one-electron operator called the Fock operator, its form is given by Equation 2.5, and $\epsilon_i$ is its eigenvalue.

$$f_i = -\frac{1}{2} \nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{j} (J_j - K_j)$$ \hspace{1cm} (2.5)

The first term corresponds to $T_e$, the second term to $V_{ne}$ and the last term describes the electronic repulsion $V_{ee}$. The two-electron operator $V_{ee}$ gives rise to the repulsive Coulomb energy ($J$), i.e. the classical Coulomb energy interaction between two charges, and the stabilizing exchange energy ($K$) that treats the Fermi correlation between electrons with the same spin. The HF equations form a set of pseudo-eigenvalue equations that are solved iteratively as the Fock operator depends on all the spin orbitals.
2. Theoretical Background

In the HF method the trial wave function consists of a single Slater determinant. For this reason, the \( V_{ee} \) term is treated as an average potential, \( v^{\text{HF}}(i) \), i.e. one electron interacts with the average potential of all others. This implies that the correlated movement of the electrons is not taken into account, i.e. the instantaneous Coulomb interactions that keep the electrons apart are not accounted for. Consequently, the energy from HF is too high – and the difference between this HF energy and the exact energy is defined as the correlation energy. It should be remembered that, while HF already determines the total energy 99% correctly, the small differences resulting from the correlation energy are relevant in chemistry, for instance in chemical reactions. Configuration interaction, Møller-Plesset Perturbation Theory and Coupled Cluster Theory are all methods that have been developed to account for the electron correlation, using the HF wave function as a starting point. The computational cost for these methods is however high and an alternative method is Density Functional Theory (DFT).

2.1.2 Density Functional Theory (DFT)

Density Functional theory (DFT) is based on the fact that the ground-state energy, \( E \) (Equation 2.6) and other properties of a system can be obtained solely from the electron density, \( \rho \), i.e. the square of the wavefunction. This fact was formulated by Hohenberg and Kohn in 1964.\(^{25}\)

\[
E[\rho(r)] = \int V_{\text{ext}}(r)\rho(r)\,dr + F[\rho(r)]
\]

(Equation 2.6)

In Equation 2.6, the first term describes the interaction between electrons and an external potential \( V_{\text{ext}} \). \( F[\rho(r)] \) describes the sum of the kinetic energy of the electrons and the inter-electronic interactions and was formulated by Kohn and Sham\(^{26}\) (Equation 2.7).

\[
F[\rho(r)] = E_{\text{KE}}[\rho(r)] + E_{\text{HI}}[\rho(r)] + E_{\text{XC}}[\rho(r)]
\]

(Equation 2.7)
2.1 Quantum Mechanics

The density is expanded in a set of one-electron orbitals, $\psi_i$, the so-called Kohn-Sham orbitals. $E_{KE}$ is the electrons' kinetic energy, $E_H$ the Coulomb energy between the electrons and $E_{XC}$ the exchange-correlation energy. $E_{XC}$ can be expressed as the sum of the exchange energy, $E_X$, which is due to correlation between electrons with the same spin (the Fermi correlation in HF theory) and the correlation energy, $E_C$, which is due to correlation between electrons with opposite spin (not treated in HF). The $E_{XC}$ term is the crucial feature in DFT, however, its exact form is not known and has to be approximated. Today, many approximate functionals are available and more still being developed.

The Local Spin Density Approximation (LSDA), based on the Thomas-Fermi model, was first formulated by Paul Dirac$^{21}$. The approximation includes a functional for the exchange energy, $E_X^{LSDA}$, which assumes that the density can be treated as a homogenous electron gas, but with different spin densities. Improvements to the LSDA method can be obtained by adding a gradient of the density and thereby considering a non-uniform electron density. Becke$^{27}$ proposed for example an improved functional for the exchange energy, $E_X^{B88}$, by adding a gradient that corrects for the wrong asymptotic behavior of $E_X^{LSDA}$.

The exact analytical expression for the correlation energy of a homogenous electron gas is not known. Quantum Monte-Carlo simulations have calculated the energy for several different densities. By interpolating the results, functionals for the uniform electron gas can be constructed. The functional for the $E_C^{VWN}$ energy was formulated by Vosko, Wilk and Nusair$^{28}$ in such a way. Lee, Yang and Parr$^{29}$ formulated a gradient corrected functional for the correlation energy, $E_C^{LYP}$, by fitting data to the helium atom.
2. Theoretical Background

2.1.3 Hybrid Hartree-Fock/Density Functional Method

Since HF fails in describing the electron correlation, but describes the exchange energy exactly, and since DFT approximates both properties, the two methods have been combined to hybrid HF/DFT. One of the most popular and widely used functionals up to date is the B3LYP functional. The B3 functional defined as Becke's 3-parameter functional and includes the exchange energies: $E_{\text{LSDA}}$, $E_{\text{HF}}$ and $\Delta E_{\text{B88}}$. The LYP functional includes the correlation energies $E_{\text{LYP}}$ and $E_{\text{VWN}}$, see Equation 2.8.

$$E_{\text{XC}}^{\text{B3LYP}} = (1-a_0)E_{\text{X}}^{\text{LSDA}} + a_0 E_{\text{X}}^{\text{HF}} + a_x \Delta E_{\text{X}}^{\text{B88}} + a_c E_{\text{C}}^{\text{LYP}} + (1-a_c)E_{\text{C}}^{\text{VWN}} \quad (2.8)$$

The constants $a_0, a_x$ and $a_c$ are parameters that have been determined by least-square fitting to experimental data for the properties of a test-set including many small molecules.

The three most important deficiencies with DFT are the self-interaction error, the near-degeneracy error and the negligence of dispersive interactions. The self-interaction error is an unphysical phenomenon due to an artificial nonzero contribution when an electron interacts with its own density. In HF the coulomb energy of the electron is directly canceled by the exchange energy. In DFT the coulomb term is described exactly while the $E_{\text{XC}}$-term is approximated leading to a non-zero difference.

The near-degeneracy error is due to the inherent single-determinant description of the wave function. For the work performed here, the consequence of this error is small since it is most common among highly charged open-shell systems.

The third limitation, the lack of van der Waals interactions, is the largest source of error. For the energies presented in Paper V, the errors in energy are addressed by adding an empirical correction. DFT-D is a method developed by Grimme and co-workers, which includes a damped atom-pairwise dispersion correction of the form $C_6 R^{-6}$.32,33
B3LYP is the most popular functional used in DFT and represents 80% of the total occurrences of density functionals in the literature 2007.\textsuperscript{34} The accuracy of this functional has therefore been exposed to numerous tests examining the accuracy with respect to various energies. Curtiss et al\textsuperscript{35} evaluated several DFT-methods on the G3/05 test set, which is an expansion of the G3/99 test set. The results show that B3LYP has a mean unsigned error of 4.1 kcal/mol. Tirado-Rives and Jorgensen tested B3LYP in respect to heats of formation and isomerization energy and found that the mean absolute errors were below 3 kcal/mol.\textsuperscript{36} The DFT-D intermolecular interaction energies were tested on the JSCH-2005 dataset\textsuperscript{37} with an error of less than 1 kcal/mol compared to high-level CCDS(T)/CBS results.\textsuperscript{38}

In this thesis all calculations were performed using the B3LYP functional on either the Gaussian 03\textsuperscript{39}, Gaussian 09\textsuperscript{40} or the JAGUAR 7.6 program\textsuperscript{41}. 
2. Theoretical Background

2.1.4 Basis Sets

The most natural way to describe the atomic orbitals is through Slater type orbitals (Equation 2.9) since these are the exact solutions of the Schrödinger equation for the hydrogen atom. However, when solving the Schrödinger equation for many-electron systems, integrals of products of the basis functions are required. These calculations can be very time consuming, so the STO:s are often replaced with Gaussian type orbitals (Equation 2.10).

\[ x^h y^k z^l e^{-\alpha r} \]  
(2.9)

\[ x^h y^k z^l e^{-\alpha r^2} \]  
(2.10)

\[ \alpha \] is a radial exponent and the sum of the Cartesian variables \( h, k \) and \( l \) determines the order of the Gaussian (for example \( h + k + l = 1 \) is a p-orbital). However, GTO:s do have disadvantages, for instance the STO has a cusp at the origin (see Figure 2.1) and the derivative at \( r=0 \) is negative while for the GTO it is zero. One way to overcome this shortcoming is to express the basis sets as a linear combination of several GTO:s (Equation 2.11) and thereby improve the description of the atomic orbitals (Figure 2.1).

\[ \phi_{\mu} = \sum_{i=1}^{L} d_{i\mu} \phi_i(\alpha_{i\mu}) \]  
(2.11)

where orbital \( \phi_{\mu} \) with the nuclear centre \( \mu \) has \( L \) functions. \( d_{i\mu} \) is the coefficient for the primitive Gaussian function \( \phi_i \), with an exponent \( \alpha_{i\mu} \). For the basis sets used in this thesis the values \( d_{i\mu} \) and \( \alpha_{i\mu} \) were predetermined and remain constant during the calculations, so-called contracted basis sets.

The so called minimal basis set corresponds to the exact number of contracted basis functions needed to describe the electronic configuration of
2.1 Quantum Mechanics

the atom(s), i.e for hydrogen and helium a single s-orbital is needed while the elements in the second row in the periodic table need two s-functions (1s and 2s) and one set of p-orbitals (1pₓ, 2pᵧ, 2pᵦ) and so forth. The basis set can be improved by doubling the basis functions – Double Zeta (DZ) type basis, i.e. two s-functions for hydrogen and helium, four s-functions and two sets of p-orbitals for the second row elements, etc.

Since chemical bonding only occurs between valence orbitals, more basis functions are often assigned to describing the valence orbitals, producing a split valence basis set. A basis set, which includes one contraction for the core orbitals and two contractions for the valence orbitals is denoted Valence Double Zeta (VDZ). Further improvements can be achieved by including more functions – Valence Triple Zeta (VTZ) and Valence Quadruple Zeta (VQZ).

Other common improvements are higher angular momentum polarization functions and diffuse functions. Higher polarization functions include p- and d-functions for hydrogen and helium and d- and f-functions for second-row atoms. Diffuse functions have small α-exponents, which are important whenever loosely bound electrons are present (for example anions). The basis sets employed in this thesis are listed in Table 2.1.

Figure 2.1 Comparison of Slater type orbitals and linear combinations of 1 to 4 Gaussian type orbitals.
2. Theoretical Background

**Table 2.1** Names of the basis set used in this thesis and their definitions

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-31G(d,p)</td>
<td>VDZ including a set of polarization functions for all atoms.\textsuperscript{44,45}</td>
</tr>
<tr>
<td>6-31++G(d,p)</td>
<td>VDZ including a set of polarization and a set of diffuse functions for all atoms.\textsuperscript{45,46}</td>
</tr>
<tr>
<td>6-311G(d,p)</td>
<td>VTZ including a set of polarization functions for all atoms.\textsuperscript{47,48}</td>
</tr>
<tr>
<td>cc-pVTZ(-f)</td>
<td>VTZ including a double set of polarization functions but not including the f-functions\textsuperscript{49}</td>
</tr>
<tr>
<td>6-311++G(d,p)</td>
<td>VTZ including a set of polarizations and a set of diffuse functions for all atoms.\textsuperscript{47,48,50}</td>
</tr>
<tr>
<td>aug-cc-pVTZ</td>
<td>VTZ including a double set of polarization functions and one set of diffuse functions for all atoms.\textsuperscript{49,51}</td>
</tr>
<tr>
<td>6-311++G(3df, 3pd)</td>
<td>VTZ including four polarization functions and one set of diffuse functions for all atoms.\textsuperscript{47,50,52}</td>
</tr>
</tbody>
</table>

2.2 Solvation Effects

Solvation effects are of great importance within chemistry and can be described theoretically in several different ways. Unfortunately, all the methods have different drawbacks and still today there is an interest in finding more suitable methods. The solvation effects can mainly be subdivided into long-range and short-range effects where the long-range effects are due to polarization and dipole orientation. The short-range can be due to hydrogen bonding, charge transfer, dispersion, hydrophobic effects, etc. In this thesis the solvation effects are often described with a continuum model, which represents the long-range effects with no explicit interactions between the solute and the solvent model. The method is computationally inexpensive compared to including hundreds or thousands of solvent molecules.
2.2 Solvation Effects

Continuum models consider the solvent as an unstructured dielectric medium or a conductor that acts as a “bulk medium”. Inside the bulk medium a cavity is placed containing the solute molecule. The solvation energy depends on three factors, the first factor is the creation of the cavity which is destabilizing and therefore requires energy. The second factor is the dispersion interactions between the solvent and the solute, which are dominated by the stabilizing van der Waals energy. The last factor is the electrostatic stabilization that occurs due to the solute’s electric charge distribution, which will polarize the medium. The two first factors are in general combined and assumed to be proportional to the surface area. The electrostatic component is in general the more important factor and can be approximated using several different methods.

The solute’s electric charge distribution polarizes the dielectric medium, which in turn induces an electric field (the reaction field) on the solute. The reaction field can be considered a perturbation to the Hamiltonian \( H_{\text{tot}} = H_0 + H_{\text{RF}} \) for which a new wave function is determined. The interaction between the solvent and solute is calculated iteratively using one of many Self-Consistent Reaction Field (SCRF)\textsuperscript{53} methods. In this thesis two different methods have been used depending on the purpose of the calculation and on the software.

In Papers I-IV small phosphate molecules were optimized in conjunction with the continuum model. For these calculation the CPCM method was used where the reaction field is introduced as polarization charge spread on the cavity surface. This method is incorporated in the Gaussian program\textsuperscript{39}.

In Paper V the Poisson-Boltzmann (PB) method was used for single-point energy calculations. Here the reaction field is introduced numerically by the PB equation solver. This method in incorporated in the JAGUAR program\textsuperscript{41}.
2. Theoretical Background

2.3 Vibrations

Infrared spectroscopy relates to the changes in vibrational energy of a molecule where the energy levels are predicted using quantum mechanics. All molecules have $3N$ degrees of freedom where $N$ is the total number of atoms. Five or six degrees of freedom, depending on whether the molecule is linear or not, are referred to the translational and rotational movements of the molecule. The remaining degrees of freedom ($3N-5$ or $3N-6$) describe the vibrations. A molecule has three types of vibrations – stretching (change of bond length), deformation (change of bond angles) and torsional vibrations, which are divided into subgroups, see Figure 2.2. A normal mode is the ensemble of all coupled vibrational movements of atoms that occur at a specific frequency. Its characteristics include that a system acts as simple harmonic oscillators where the atoms' center of mass do not move and where all atoms are at their equilibrium positions simultaneously.

![Figure 2.2](image)

**Figure 2.2 The different types of vibrations where the “+” and “-” are out of the plane vibrations. The main features of this picture are taken from Hesse et al.**

In classical mechanics the vibrational mode for a diatomic molecule can
be described as in Figure 2.3. This model consists of two point masses, $m_1$ and $m_2$, which represent the atoms. The atoms are joined together by an elastic spring with a force constant $k$. The force constant is a measure of the bond strength between the atoms. If the vibration is derived from the harmonic oscillator, the potential energy ($V$) will depend on the internuclear distance $r$ and the force constant (Equation 2.12).

$$V(r) = \frac{1}{2} k \cdot (r-r_0)^2 \quad (2.12)$$

$r$ is the actual bond length and $r_0$ is the equilibrium bond length where the potential energy is the lowest. The potential energy can be expanded in a Taylor series centered at $r_0$ and where $V_e$ is the potential energy at equilibrium, which can be defined as zero, see Equation 2.13. The second term in Equation 2.13 is the slope of the potential energy at the equilibrium and is equal to zero. By neglecting cubic and higher order terms, the potential energy surface can be described as a quadratic function of $r$ – this is known as the harmonic approximation. The higher order terms serve as anharmonic corrections to the potential. From Equations 2.12 and 2.13 an expression for the force constant can be derived by taking the second derivative of the potential energy with respect to $r$.

$$V = V_e + \left( \frac{\partial V}{\partial r} \right)_e (r-r_0) + \frac{1}{2} \left( \frac{\partial^2 V}{\partial r^2} \right)_e (r-r_0)^2 \quad (2.13)$$

In quantum mechanics, the solution to the Schrödinger equation for a diatomic molecule undergoing harmonic motion is given by the vibrational energy $E_n$ with the vibrational quantum number $n$ (Equation 2.14).
2. Theoretical Background

\[ E_n = \left( n + \frac{1}{2} \right) \cdot h \cdot \nu \quad n=0, 1, 2, ... \]  

(2.14)

h is Planck's constant and \( \nu \) the vibrational frequency which depends on the force constant \( k \) and the reduced mass \( \mu \) (Equation 2.15).

\[ \nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \]  

where \( \mu = \frac{m_1 \cdot m_2}{m_1 + m_2} \)  

(2.15)

For polyatomic systems, the potential energy is approximated as a quadratic function of its internal coordinates. The Harmonic approximation thereby approximates each non-linear molecule as though it consists of 3N-6 harmonic oscillators. The second derivatives result in the Hessian matrix where the eigenvalues are \( 4\pi^2 \mu \nu^2 \) and yield the vibrational frequencies \( \nu \). The eigenvectors are used to describe the normal modes, the absorbance and the atoms' displacements. In this thesis the frequencies are given in wavenumbers, which are calculated by dividing the frequencies with speed of light.

2.4 Natural Bond Orbital Theory

Molecular orbitals describe the wave-like behavior of an electron. Natural bond orbitals (NBOs) are the Lewis structure representation of the wave function and are calculated in the NBO 3.1 software\textsuperscript{55} incorporated in the GAUSSIAN 03 program\textsuperscript{39}. NBOs\textsuperscript{56} are a set of orthonormal orbitals with maximum occupancy (highest percentage of the total electron density) which gives the most accurate Lewis-like structure. The orbitals consist of localized electron pair orbitals that describe bonds and lone pairs. Departures from a single localized Lewis structure are called “non-Lewis-type” NBOs and they describe delocalization effects.

Negative hyperconjugation involves the interaction between a filled n-orbital (donor) and an empty antibonding \( \sigma^* \)-orbital (acceptor), which produces an extended molecular orbital. Mixing between filled orbitals and antibonding orbitals describes departures of the electron distribution from an
idealized Lewis structure. Negative hyperconjugation lowers the total energy of the molecule and destabilizes the σ-bond. In this thesis the interaction between the phosphates terminal oxygens’ n-orbitals (donors) and the σ*-orbital of the phosphorus and the bridging oxygen (acceptor) is studied, producing an elongated σ-bond between the phosphorus and the bridging oxygen, see Figure 2.4. The stabilization energy, E(2), estimates the strength of the interaction and is calculated through second order perturbation theory analysis of the Fock matrix, $F_{ij}$, which describes the donor-acceptor interaction (Equation 2.16).

$$E(2) = -q_{\text{donor}} \frac{F_{ij}^2}{\Delta E}$$  \hspace{1cm} (2.16)$$

$\Delta E$ is the energy gap between the two orbitals and $q_{\text{donor}}$ is the occupancy of the donor orbital.

**Figure 3.4** The figure shows the overlap (in stripes) between the natural bond orbitals of $\sigma^*(P-O_B)$ and $n(O_T)$ for AcP$^{2-}$. 
“The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity.”

Albert Einstein (1879-1955)
3. Benchmark Studies (Paper II)

3.1 Introduction

In the past decades the progress of applying quantum mechanics to chemistry has been extraordinary. The continuous development of computational power and the development of hybrid DFT the past 20 years have made theoretical calculations routine among scientists and today quantum chemistry is not only used as a complement to experiment but a tool of its own. DFT does have its limitations however and it is therefore important to perform benchmark tests.

This chapter will mainly focus on the results from Paper II a basis set and solvation effect study on vibrational wavenumbers. The calculations were performed on the molecules in Figure 3.1.

![Methyl Phosphate (MP$^3$)](image)
![Dimethyl Phosphate (DMP$^*$, HDMP)](image)
![Acetyl Phosphate (AcP$^5$)](image)
![p-Tolyl Phosphate (TP$^5$)](image)
![Phosphoenol Pyruvate (PEP$^5$)](image)

Figure 3.1 *The molecules studied and their abbreviations.*
3.1 Introduction

3.2 Basis Set Effects on Wavenumbers

Today there exist many basis set benchmarking studies based on energies. The basis set study presented here is based on the wavenumbers of the phosphate groups shown in Figure 3.1. The mean absolute error (MAE) and mean signed error (MSE) were investigated comparing the wavenumbers of the stretching P-O vibrations to either experimental results or to the wavenumbers calculated with the largest basis set for several different basis set, see Table 3.1. The results show that the 6-31G basis set has the largest absolute error (>100 cm⁻¹), that the 6-31G(d,p), 6-31++G(d,p), 6-311+ +G(d,p) and 6-311G(3df, 3pd) basis sets all have an absolute error between 20 and 30 cm⁻¹, that the 6-311G** and the aug-cc-pVTZ basis sets have a MAE close to 15 cm⁻¹ while the two largest basis sets have a MAE less than 10 cm⁻¹. Why the MAE is only around 15 cm⁻¹ for the 6-311G** basis set is not known but most likely due to cancelation of errors.

The MSE in Table 2 show that the polarization functions upshift the wavenumbers while the diffuse functions downshift the wavenumbers. The effect of the extra valence function is more obscure. The best results are obtained with the 6-311++G(3df, 3pd) basis set, however it is also the most time consuming. The 6-31++G** basis set has been a good choice in balancing computational time with good results. The wavenumbers calculated with this basis set are more consistent than the wavenumbers calculated with either the 6-31G(d,p) or the 6-311G** basis sets. The 6-31+ +G** basis set is also more affordable than the larger basis sets. In Papers I, III and IV the majority of the work was therefore calculated with the 6-31+ +G** basis set.

The calculations performed in Paper V do not include studies of small phosphate molecules nor on frequencies. The justification for the basis set in that paper is therefore based on results from the literature. In Paper II it is concluded that the geometries calculated with the 6-31G** basis set are
3.2 Basis Set Effects on Wavenumbers

quite close to the geometries optimized with the larger basis sets. This was also concluded by Siegbahn\textsuperscript{57} who investigated the different steps of the manganese catalase using three different basis sets. Siegbahn also concludes that the 6-31G basis set can be used in geometry optimizations as long as the relative energies are calculated with larger basis sets (single point). In Paper V, the geometry optimizations were performed with the 6-31G** basis set and single-point energy calculations were performed with the cc-pVTZ basis set.

Table 3.1 The mean absolute error (MAE) and mean signed error (MSE) for different basis sets when calculating the wavenumbers of the P-O stretching vibrations of the molecules in Figure 3.1. Either experimental wavenumbers (exp) or wavenumbers calculated with B3LYP/6-311++G(3df, 3pd) (calc) are used as a reference.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>MAE-exp (cm(^{-1}))</th>
<th>MAE-calc (cm(^{-1}))</th>
<th>MSE-calc (cm(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-31G</td>
<td>143</td>
<td>122</td>
<td>-122</td>
</tr>
<tr>
<td>6-31G**</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>6-31++G**</td>
<td>28</td>
<td>22</td>
<td>-20</td>
</tr>
<tr>
<td>6-311G**</td>
<td>16</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>6-311++G**</td>
<td>29</td>
<td>25</td>
<td>-22</td>
</tr>
<tr>
<td>6-311G(3df, 3pd)</td>
<td>30</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>aug-cc-pVTZ</td>
<td>15</td>
<td>10</td>
<td>-10</td>
</tr>
<tr>
<td>6-31++G(3df, 3pd)</td>
<td>9</td>
<td>5</td>
<td>-3</td>
</tr>
<tr>
<td>6-311++G(3df, 3pd)</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Benchmark Studies (Paper II)

3.3 Solvation Effects on Wavenumbers

The majority of all chemical processes are influenced by its surrounding either directly or indirectly. In the first case, explicit solvent molecules need to be included in order to model the effect while in an indirect case it may be enough to include a continuum model.

The effect of the CPCM-continuum model on the wavenumbers of the phosphorylated molecules in Figure 3 was studied. The MAE in wavenumbers when comparing gas phase calculations with calculations optimized using the CPCM-continuum model was 33 cm\(^{-1}\). All the geometries were optimized with the 6-311++G(3df, 3pd) basis set. The results also show that the MAE partly depends on the type of vibration: the error of the antisymmetric P-O\(^\tau\) stretching vibration of AcP\(^2-\) was 77 cm\(^{-1}\) while the symmetric stretch for the same molecule and bond was 1 cm\(^{-1}\). [Paper II]

Explicit water molecules are necessary in order to calculate the vibrational wavenumbers of protonated groups, such as POH and COH. Without the explicit water molecules, the wavenumbers are significantly underestimated, as is seen for both singly and doubly protonated PEP. We studied this effect by adding four explicit water molecules to the protonated PEP molecules and found that the wavenumbers of the P-O(H) and C-O(H) groups were upshifted by 50-100 cm\(^{-1}\) and much closer to experimental results (but still underestimated). The remaining vibrations hardly shifted at all. [Paper I]

The effect of explicit interactions were studied in more detail in Paper III and IV. The results show how different interactions affect the geometry and the wavenumbers of the phosphate molecules. These results are presented in the next chapter.
“Everything has its beauty but not everyone sees it.”

Confucius (551-479 BC)
4. The Infrared Spectra of Phosphoenol Pyruvate (Paper I)

In the final step of the glycolysis, pyruvate is formed from phosphoenol pyruvate (PEP) in a reaction catalyzed by pyruvate-kinase (PK). The structural changes of PK due to the binding of PEP can be studied using infrared spectroscopy. However, assigning the observed absorption bands to vibrational modes is not always simple. Knowledge about group wavenumbers, a strategy that does not take possible vibrational couplings into account, and isotope-labeling are two strategies that are commonly used. A third common method which has become more and more popular is to compare the experiments with calculations and thereby assign the different absorption bands.

The chemical reactivity of PEP has been studied theoretically before in terms of the hard-soft acid-base principle in enzymatic catalysis. The study indicated that enzymes can control the reactivity by conformational changes of PEP, specifically through the dihedral angle between the carboxylate plane and the enol plane.58

In this chapter the assignments of the experimental infrared spectra of PEP in three different protonation states and theoretical frequency shifts due to the rotation of the dihedral angle O-C1-C2-O of fully-ionized PEP will be presented. These are two of the main results from Paper I, results which have been applied in new infrared studies of PEP binding to PK2,3,59.
4. The Infrared Spectra of Phosphoenol Pyruvate (Paper I)

4.1 Normal Mode Assignment of PEP

Depending on the pH, PEP exists in three different ionization states, see Figure 4.1. At lower pH-values, the protonated PEP can adopt different conformations: a linear conformation or those that include internal hydrogen-bonds. When assigning the experimental absorption bands, all conformations need to be regarded. At all ionization states, PEP can adopt the linear conformation. At pH 4.6, only the phosphate group is protonated, the singly-protonated (SP) PEP can therefore also adopt the conformation with the hydrogen-bond from the phosphate group to the carboxyl group. At pH 2.1, both carboxyl- and phosphate groups are protonated. Doubly-protonated (DP) PEP can therefore adopt all three heavy-atom conformations shown in Figure 4.1.

![Figure 4.1](image)

**Figure 4.1** Three different heavy-atom conformations of PEP: a) The linear conformation, b) hydrogen bond from the phosphate group to the carboxyl group and c) hydrogen bond from the carboxyl group to the phosphate group.
4.1 Normal Mode Assignment of PEP

The calculated wavenumbers of PEP are very close to the experimental results (see Figure 4.2) and a complete band assignment has been achieved in the spectral range from 1800 to 800 cm\(^{-1}\). Exceptions are the vibrations involving the hydroxyl groups for the protonated PEP. The results show that hydroxyl groups are very sensitive to hydrogen bonding and are therefore significantly underestimated (see Chapter 2) when using a continuum model for water.

The calculations show that many of the experimental normal modes are due to the vibration of predominantly one internal coordinate. However, significant coupling occurs between the CH\(_2\)-group deformation mode and with either the symmetric stretching vibration of COO\(^{-}\) or the stretching vibration of the C-O(P) bond.

![Figure 4.2](image-url) *Figure 4.2 The infrared spectrum of fully-ionized PEP. The top panel shows the experimental spectrum at pH 9 and the lower panel shows the calculated spectrum. The vibrations were analysed by calculating the potential energy distribution (PED) for each normal mode.*
4.2 Reactivity of PEP from IR-spectroscopy

Li and Evans\textsuperscript{58} have shown that the O-C-C-O dihedral angle size is correlated with the reactivity of PEP in enzymes. We wanted to determine if a widening of the dihedral angle can be observed by IR. The calculated IR spectrum of PEP\textsuperscript{3-} was therefore studied by rotating the O-C-C-O dihedral angle 30, 50, 70 and 90°, see Figure 4.3. For these calculations, wavenumber shifts of 10-40 cm\textsuperscript{-1} were observed for the two modes that contributed to the experimental band at 1410 cm\textsuperscript{-1} (stretching vibrations of the carboxylic- and the methyl groups). The shifts of the 1410 cm\textsuperscript{-1} band indicate that the reactivity can be assessed by IR spectroscopy, which thereby can provide important insights of the enzymatic mechanism.

\textbf{Figure 4.3} The calculated IR-spectrum of FI-PEP during rotation of the dihedral angle O-C-C-O. The dihedral angle is frozen for all the calculations except for when the angle is 0°. The figure also shows the relation between the two modes of the 1410 cm\textsuperscript{-1} band. As the O-C-C-O dihedral angle increases, the band gap increases.
"In science one tries to tell people, in such a way as to be understood by everyone, something that no one ever knew before. But in poetry, it's the exact opposite."

Paul Dirac (1902-1984)
5. The Bitter Devil of E2P (Papers III-V)

The hydrolysis of the phosphorylated Asp351 of E2P in Ca\textsuperscript{2+}-ATPase, is \(10^6\)-\(10^7\)-fold faster compared to the model compound acetyl phosphate in water. According to the classical transition state theory, the rate constant can be expressed in terms of Gibbs free energy of the activation barrier, \(\Delta G^\ddagger\), and is determined by Eyrings equation, see Equation 5.1.\textsuperscript{60}

\[
  k = \frac{k_B T}{h} e^{\frac{\Delta G^\ddagger}{RT}}
\]  

(5.1)

\(k_B\) is Boltzmann’s constant and \(h\) is Planck’s constant. At room temperature (298.15 K), the \(10^6\)-\(10^7\)-fold rate enhancement thereby corresponds to a lowering of the activation barrier with 8-9 kcal/mol. This lowering in energy is an essential feature of the Ca\textsuperscript{2+}-ATPase since it is required for fast muscle relaxation.

The rate enhancement of many enzymes is due to specific interactions between the protein and the substrate, which are commonly known to stabilize the transition state. The goal of Paper III was to gain information about the molecular environment in order to better understand the importance of the specific interactions. Several simple models were therefore studied, including one of three phosphorylated molecules in its dianionic form: acetyl phosphate, methyl phosphate or p-tolyl phosphate. Explicit HF and H\textsubscript{2}O molecules were added in order to model environmental effects and interactions that may occur in a protein matrix. The structure and partial charges of the models were compared to more realistic models of the active site of E2P.

The underlying catalytic mechanism of E2P was earlier studied within our
5. The Bitter Devil of E2P (Papers III-V)

group by characterizing the vibrational spectrum. An isotope exchange experiment of the E2P phosphate was then evaluated with empirical correlations. The aim of Paper IV was to use the same models as in Paper III and evaluate the experimental results theoretically.

In Paper V, the energy-profile of the hydrolysis was investigated. This can be seen as the final step in explaining the rate-enhancement of the Ca\textsuperscript{2+}-ATPase at an atomistic level.

5.1 Simple environment models

Methyl phosphate (MP), p-tolyl phosphate (TP) and acetyl phosphate (AcP) are representative models for aspartyl phosphate, serine or threonine phosphate, and tyrosine phosphate, respectively. In Papers III and IV, the molecules interacted explicitly with either HF or H\textsubscript{2}O molecules. The HF molecule was chosen due to its high polarity leading to strong electrostatic interactions and H\textsubscript{2}O since it is a physiological molecule.

Since water energetically prefers to form hydrogen bonds to more than one of phosphate’s oxygen atoms, it is simpler to study the effect of single interactions with HF-molecules. Such models included hydrogen bonds from HF to the terminal oxygens of the phosphate (O\textsubscript{T}), to the bridging oxygen (O\textsubscript{B}), and for AcP, to the carbonyl oxygen, (C=)O, see Figure 5.1. The interaction between the phosphates' phosphorus atom and the fluorine atom of HF was also studied.
5.1 Simple environment models

![Figure 5.1](image.png)

**Figure 5.1** The interactions between the phosphorylated molecules and HF. The HF molecule was placed at several fixed distances from the O_T, O_B, P, and (C=)O atoms.

Interactions with water being more interesting in a biological perspective, the interactions between the phosphorylated molecules and water were also considered. All in all over 150 models were examined: over 60 included one HF-molecule, over 60 included two or more HF-molecules and over 30 were modeled with three or more H_2O-molecules. In all but eight models, the interacting molecules were fixed to a specific distance from one of the O_T, O_B, P, and (C=)O atoms. The geometry of the models were optimized with the constraints, using the B3LYP functional and three different basis sets depending on the purpose of the calculation; 6-31G** (referred to as small), 6-31++G** and 6-311++G(3df, 3pd) (referred to as large). For the majority of the calculations, the 6-31++G** basis set was used, and unless mentioned in the text, this is the basis set used. For all the geometry optimizations, solvent effects were taken into account with the solvation model, CPCM, and with an ϵ of either 4 (modeling a protein environment^{61,62}) or 78.39(modeling water).

The aim for one of the simple amino acid models in Paper IV, was to construct a model that reproduced the experimental wavenumbers of the phosphate group of E2P (1194 and 1137 cm⁻¹). The constraints for this model were chosen after examining the interactions found between the phosphorylated Asp351 and its protein environment (see Figure 5.2). The
5. The Bitter Devil of E2P (Papers III-V)

The exact wavenumbers (1194 and 1137 cm\(^{-1}\)) were reproduced using the model in Figure 5.2, with the large basis set.

**Figure 5.2** The inactive model (see Section 5.2) of E2P and a simple amino acid model. This specific simple model reproduces the experimental wavenumbers of the P-O stretching vibration of E2P (1194 and 1137 cm\(^{-1}\)). 1.80 Å, 2.00 Å and 1.50 Å are the three constraints. No listed R-values were constrained.
5.2 Protein environment models

The Ca\(^{2+}\)-ATPase of the skeletal muscle sarcoplasmic reticulum (SERCA1a) is structurally and functionally the best studied member of the P-type ATPase family.\(^{64}\) The crystal structure has been solved for several different states in the reaction cycle.\(^{7,8,65-68}\) The crystal structure of the E2P intermediate alone exists in several different states, however all structures include one of three phosphate analogues: BF\(_3\), AlF\(_4\) or MgF\(_4\). The three different analogues correspond to three different states that occur during the hydrolysis reaction. E2•BeF corresponds to the E2P ground-state, the E2•AlF to the transition-state and E2•AlF to the product of the dephosphorylation reaction. The active site of E2•AlF and E2•MgF are very similar and are therefore unitedly regarded as the “active state” in this thesis while the E2•BeF will from now on be regarded as the “inactive state”. The largest difference between the two states is the positioning of the Glu183-loop. In the “inactive state”, Glu183 is pointing away from the active site and the space around the BeF\(_3\) molecule is very compact, leaving no space for an attacking water molecule. In the “active state” Glu183 is in close proximity to the phosphate analogues. The function of the Glu183 is to position an attacking water molecule close to the reactant and thereafter to serve as a base during the dephosphorylation of Asp351.\(^{5,64}\) In Papers III and IV different properties of the phosphorylated Asp351-residue were studied while in Paper V the reaction mechanism of the dephosphorylation was determined.

The active site for the two states were modeled using the quantum chemical cluster approach.\(^{62}\) The aim was to perform accurate QM calculations on a small but carefully selected fragment of the protein. The active site of E2-P was constructed from two different X-ray structures of the E2P analog (PDB: 1WPG or 2ZBF) where the phosphate analogs, MgF\(_4\)/BF\(_3\), were substituted for phosphate. The residues and the water molecules that were expected to be of importance were extracted from the PDB-file together with the metal ion. The model, see Figure 5.3, included Asp351 bound to phosphate, a magnesium ion, water molecules (two of
5. The Bitter Devil of E2P (Papers III-V)

which were ligated to Mg$^{2+}$) and the following residues; Thr181, Gly182, Glu183, Thr353, Thr625, Gly626, Lys684, Asp703, Asn706, and Asp707. To further reduce the model size, the residues were truncated so that only the side chains and a few important peptide backbones were included. For the peptide links that were not considered to be involved, the peptide’s carboxyl carbon and nitrogen were substituted for hydrogen atoms and fixed to the X-ray coordinates along with the α-carbon atoms. This is a common procedure to keep the backbone positioned and to prevent unnatural movements thereby keeping the optimized structure close to the crystal structure. The hydrogens were added manually. The side chains of glutamate, aspartate and lysine are all charged, which is in accordance with the physiological pH at 7.4. The total charge of the system was -2, which is the charge of the phosphate substrate, and is in accordance with the model studied by Himo et al who studied a similar system.

In Paper III, all models used were optimized using the 6-31++G** basis set. Since it is not computationally affordable to perform frequency calculations using this basis set on such large systems, the smaller 6-31G** was used for both the geometry optimization and the frequency calculations in Papers IV and V. The energies obtained from this smaller basis set in Paper V were refined by including single-point calculated energies with the larger basis set cc-pVTZ(-f) (6-311G* for Mg$^{2+}$). Both zero-point corrections from the frequency calculations and dispersion corrections according to Grimmes DFT-D method, which is implemented in the XYZ-viewer, were added to the total energy. The effects of the protein environment were estimated by the solvation energy, included by performing single-point calculations with a homogenous polarizable medium on the gas-phase, optimized structure. The choice of dielectric constant is of less importance since systematic studies have shown that the solvation effects saturate when the model reaches a size of 150-200 atoms. In Paper V, ε=4 was used however since it is considered to represent the protein environment well.

$\Delta G^\ddagger$ is approximated with the enthalpy in this thesis. Since several atoms are locked into specific coordinates, accurate entropy cannot be calculated.
5.2 Protein environment models

However, the effect of the entropy is expected to be small.

![Image](image_url)

**Figure 5.3** The Ca$^{2+}$-ATPase (PDB: 1WPG) and a model of its active site. In the model the protein environment is represented by continuum with a dielectric constant of 4. The fixed atoms are encircled in red.

5.3 Structural and Energetic Effects (Paper III)

Enzymes work by lowering the activation energy of a reaction and thereby increasing the reaction rate. This can be achieved by stabilizing the transition state and the most effective way to do so is through electrostatic interactions. The same interactions can stabilize the ground state (GS) to a configuration that is more or less transition-state like. For the Ca$^{2+}$-ATPase, a more TS-like reactant in the GS could help to explain the faster hydrolysis of the phosphorylated Asp351, as compared to AcP in water. By studying electrostatic interactions on simple phosphorylated amino acid models, both structurally and energetically, the effect of these interactions can be evaluated.

Energetically, specific interactions were studied where the phosphorylated amino acid models (see Section 5.1) interacted with both HF and H$_2$O. The O$_T$⋯H interactions resulted in the energetically most stable conformations,
while the least stabilizing effect was observed in the O_B···H interactions.
Between the phosphorylated molecules, the hydrogen bonds stabilized MP the most and AcP the least.

While close hydrogen-bonds to the O_B atom and to the (C=)O atom have an energetically stabilizing effect, the calculations show that they have a destabilizing effect on the P-O_B bond. In other words, an approach of the HF molecule to the O_B atom will elongate the P-O_B bond. This effect is much stronger for an O_B···H interaction compared to a (C=)O interaction. For the O_T···H interactions, the effect is opposite: an energetically stabilizing effect will shorten the P-O_B bond length.

Two GS models of the E2P intermediate of the Ca^{2+}-ATPase were studied; the “active model” and the “inactive model”. The largest difference between these two models is the positioning of the TGES-loop as discussed in section 5.2. Another difference is the positioning of a hydrogen-bond between the NH_3-group of Lys684 and the phosphorylated Asp351. The hydrogen-bond shifts from interacting with the phosphate's O_T atom in the “inactive model” to interacting with the phosphate's O_B atom in the “active model”. From our conclusions above, such a shift in interaction should destabilize the “inactive” ground-state compared to the “active state” both energetically (the O_T···H interaction is energetically favorable compared to O_B···H) and structurally (the P-O_B bond length). Energetically, the two models can not be righteously compared since the fixed atoms are not constrained into the same coordinates and due to entropic effects. Structurally, the P-O_B bond length of the phosphate molecule is 0.04 Å longer in the “active model” compared to the “inactive model”.

5.4 Charge (Paper III)

In the previous section, different electrostatic interactions were described as well as how they differed in the “active model” and the “inactive model”. After the repositioning of the TGES-loop, the H-bond from Lys684 shifts
5.4 Charge (Paper III)

from interacting with $O_T$ to $O_B$ of the phosphorylated Asp351, which can be assumed to destabilize the ground-state. Electrostatic interactions are due to the non-uniform distribution of positive and negative charge. In a protein, these interactions reflect a preorganized polar environment of the enzymes active site.\textsuperscript{76} In the following section, charge changes induced by the interactions between HF/H2O and the phosphorylated molecules are described, thereby showing how an environment can stabilize the dianionic phosphate group through electrostatic interactions.

There are two ways for the dianionic phosphate group to spread its charge, it can either transfer some of the charge to the interacting molecules, or the P-O$_B$ bond can be elongated and thereby transfer charge to $O_B$. The charge of the phosphate molecule, that of the interacting molecules and that of the whole organic group (O$_B$-R) was, i.e. subtracting the charge of the phosphate group and the interacting molecules from the total charge of the model. Since the charge of the O$_B$ atom can be delocalized to the organic group that it is bonded to, it is more appropriate to study this charge than the charge of only the O$_B$ atom.

First, the charge effect of single interactions was studied between the three phosphorylated molecules and one HF-molecule. For all the models with a hydrogen bond interaction (O···H), the charge of the HF molecule decreased and became more negatively charged when the O···H distance was shortened. For all interactions, the negative charge of the phosphate group decreased, when the interacting molecule approached. For the organic group, the effect of the interactions differed; shorter O$_T$···H interactions led to a less negative charge since the P-O$_B$ bond was shortened. During an approach of the HF to the O$_B$ atom, the charge increased (more negative), which is explained by a longer P-O$_B$ bond. However, for very short O$_B$···H interactions, the charge of the organic group started to decrease since the length of the P-O$_B$ bond was no longer affected.

In summary, all interactions with the phosphate molecules lead to a less negative charge, while the effect of the interactions on the organic group is more complex. The charge of the organic groups seem to be primarily
5. The Bitter Devil of E2P (Papers III-V)

affected by the P-O\textsubscript{B} bond length, and not by charge transfer to the interacting molecule. By studying all the simple amino acid models, a correlation was found between the charge of the organic group and the P-O\textsubscript{B} bond length.

In the Ca\textsuperscript{2+}-ATPase, the H-bond from the NH\textsubscript{3} group of Lys684 shifted from interacting with O\textsubscript{T} in the “inactive model”, to interacting with O\textsubscript{B} in the “active model”. From the simple models, such a shift contributed to a more negatively charged O\textsubscript{B} atom and a longer P-O\textsubscript{B} bond. Justly, the P-O\textsubscript{B} bond of the “active model” is 0.04 Å longer and the charge of the O\textsubscript{B} atom is 0.03e lower compared to the “inactive model”.

5.5 Vibrations (Paper IV)

Infrared spectroscopy is an excellent method for investigating molecular reactions at an atomic level. Due to its high sensitivity, which surpasses the resolution of X-ray crystallography, information on catalytic properties can be extracted. However, despite the great potential of infrared spectroscopy, it is easy to misinterpret the spectrum. Quantum mechanical calculations are therefore an optimal complement to infrared spectroscopy. Here below, the results from both the simplified phosphorylated amino acid models and the more realistic models are presented, as well as how the theoretical work can be an appropriate complement to infrared spectroscopy.

It has been shown that the P-O bond order and bond length for phosphate groups can be determined from vibrational spectroscopy using the bond valence sum derived by Brown and Wu.\textsuperscript{77,78} This empirical method relates the fundamental stretching frequency to both the bond order, and the bond length. For phosphate molecules, the wavenumber of the fundamental stretching frequency is calculated using the root mean square of the molecules’ three wavenumbers, These wavenumbers are a result of the P-O\textsubscript{T} stretching (two asymmetric and one symmetric) vibrations.
The wavenumbers of the asymmetric and symmetric P-O\textsubscript{T} stretching vibrations were calculated for the simplified phosphorylated amino acid. Paper IV showed that the symmetric stretching vibration, $v_s$, of the phosphate group does not depend on the P-O\textsubscript{B} bond length (Figure 5). However, the correlation between the wavenumber of the average $v_{as}$(P-O) vibration and the P-O\textsubscript{B} bond length is evident (Figure 5.4). These results suggest that the bond valence sum can be improved by simply relating the bond order and bond length to the asymmetric stretching vibration, $v_{as}$, of the phosphate group, and not to that of the symmetric stretching vibration.

All intermediates of the Ca\textsuperscript{2+}-ATPase reaction cycle have previously been resolved using IR-spectroscopy.\footnote{For the intermediate E2P, the bond valence sum was then applied in order to determine the bond length of P-O\textsubscript{B} in E2P.\footnote{For these calculations, three bands at 1194, 1137 and 1115 cm\textsuperscript{-1} were assumed to stem from E2P. However, in Paper IV, none of the symmetric vibrations were calculated to be above 1000 cm\textsuperscript{-1} (Figure 5.4). The interpretation of the spectrum has therefore been modified and the band at 1115 cm\textsuperscript{-1} is now assigned to the asymmetric stretching vibration of another state. As no symmetric stretching vibration is present, the bond valence sum cannot be used to calculate this bond distance.}} For the intermediate E2P, the bond valence sum was then applied in order to determine the bond length of P-O\textsubscript{B} in E2P.\footnote{For these calculations, three bands at 1194, 1137 and 1115 cm\textsuperscript{-1} were assumed to stem from E2P. However, in Paper IV, none of the symmetric vibrations were calculated to be above 1000 cm\textsuperscript{-1} (Figure 5.4). The interpretation of the spectrum has therefore been modified and the band at 1115 cm\textsuperscript{-1} is now assigned to the asymmetric stretching vibration of another state. As no symmetric stretching vibration is present, the bond valence sum cannot be used to calculate this bond distance.}
Two interpretations of the experimental spectrum that might explain the 1115 cm\(^{-1}\) band are that the observed spectrum both represents the “inactive” and the “active states” of E2P or that it represents the ground state of both E2P and E1P. The wavenumbers for both the “inactive” and the “active model” were therefore calculated in Paper IV. The results showed that the experimental bands of 1194 and 1137 cm\(^{-1}\) were closer to the wavenumbers of the “inactive model” compared to those of the “active model”. Furthermore, none of the models applied could explain the 1115 cm\(^{-1}\) band. The calculations were performed using the small 6-31G** basis set, as it is not feasible to calculate the wavenumbers using the “standard” 6-31++G** basis set on so many atoms. In Paper II, however, it was shown that the wavenumbers of this small basis set are not always very accurate (See Chapter 3). For this reason, additional calculations were performed using a simple amino acid model as a complement.

The wavenumbers of the model shown in Figure 5.2 (righthand panel) were calculated using the large 6-311++G(3df, 3pd) basis set. This model was applied since it reproduces the exact wavenumbers of the experimental infrared spectrum (Table 5.1). The aim of this approach was to compare the wavenumbers of the simple model to both the “inactive” and the “active model”. To do so the geometry of the simple model was reoptimized using the same constraints but with the small basis set. The resulting calculated wavenumbers from this new model were seen to be closer to those of the “inactive model” than to those of the “active model” (Table 5.1). Since the wavenumbers of the “inactive state” were closer to both the experimental bands and the calculated wavenumbers of the simpler model, it can be assumed that the “inactive state” is the state to which the experimental spectrum corresponds. As the wavenumbers of the “active state” did not present an explanation to the band at 1115 cm\(^{-1}\), it can also be assumed that the experimental spectrum is unable to record the “active state”, and thus that this state is a very short-lived state. This indicates that the band at 1115 cm\(^{-1}\) is due to E1P.
5.5 Vibrations (Paper IV)

Table 5.1 The wavenumber of the antisymmetric stretching vibration of P-OT.

<table>
<thead>
<tr>
<th></th>
<th>Experiment</th>
<th>Acp large basis set</th>
<th>Acp small basis set</th>
<th>Protein model small basis set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>v_{as}(PO_T) cm^{-1}</td>
<td>v_{as}(PO_T) P-O_B cm^{-1} Å</td>
<td>v_{as}(PO_T) P-O_B cm^{-1} Å</td>
<td>v_{as}(PO_T) P-O_B cm^{-1} Å</td>
</tr>
<tr>
<td>E2P “inactive”</td>
<td>1137(^a) 1194(^b)</td>
<td>1137(^b) 1194(^b) 1.77</td>
<td>1145(^b) 1217(^b) 1.79</td>
<td>1133 1223 1.78</td>
</tr>
<tr>
<td>“active”</td>
<td></td>
<td></td>
<td></td>
<td>1131 1235 1.82</td>
</tr>
</tbody>
</table>

\(^a\)Barth and Bezlyepkina\(^76\)
\(^b\)Modeled with 3HF, see Figure 5.2

5.6 Reaction Mechanism (Paper V)

The final step to fully elucidating the fast hydrolysis of E2P is to study the actual reaction path. Applying the “active model” (Section 5.2) as the reactant, the energy profile for the hydrolysis was studied. In the central part of the model, the phosphate group is bound to Asp351 (R1=1.82 Å) and is separated from the nucleophile, an attacking water molecule that is referred to as water1, by 4.08 Å (R2). Water1 is kept in place by a hydrogen bond to the residue Glu183 (R3 = 1.73 Å).

Approximate transition states (TS) were obtained by optimizing the structure for different constrained distances of O-P and a full TS-optimization was performed starting from the structure with the highest energy. The vibrational frequencies of this TS were analyzed to ensure that the main imaginary frequency corresponded to the correct motion; i.e. that the bond between the phosphate group and Asp351 was broken and that the bond between the phosphate group and water1 was formed. The energy barrier was calculated to be 14.3 kcal/mol, which is in agreement with the experimental rate of 0.1 – 115 s\(^{-1}\) (depending on pH, temperature and buffer), i.e. an activation barrier of 15-19 kcal/mol.\(^79\)\(^-\)\(^84\) This is further supported by the results presented in Sections 5.3 and 5.5, which indicated that the “inactive model” corresponds to the true ground-state, and further
5. The Bitter Devil of E2P (Papers III-V)

that an extra energy most likely needs to be added due to the TGES-loop reorientation. This would thereby increase the activation barrier with a few kcal/mol, and thus correlate even better with experiments.

The product, a single protonated phosphate molecule, was calculated to be almost thermoneutral (+0.9 kcal/mol). The calculations showed that the Glu183 residue acts as a base, accepting a proton from water1, and thereby activating the attacking water molecule. In water, such a base is not found and a main conclusion is therefore that the Glu183 is partly responsible for the fast hydrolysis of phosphorylated Asp351 in E2P, compared to acetyl phosphate in water. The calculated energies for the reaction step showed that the reaction is reversible, which is in agreement with the reaction cycle of Ca\(^{2+}\)-ATPase, where all steps are reversible. After the hydrolysis, the phosphate group is released and according to Pickart and Jencks\(^79\), this release results in a ~2 kcal/mol downhill reaction compared to the ground-state, which would make the hydrolysis reaction exothermic.
“The beauty of a living thing is not the atoms that go into it, but the way those atoms are put together.”

Carl Sagan (1934-1996)
6. Conclusions and Outlook

The results are read and you can now relax. The Ca\textsuperscript{2+}-ATPase is working at its fullest, consuming ATP. The hydrolysis reaction of E2P is fast thanks to Glu183 acting as a base and the electrostatic interactions found at the active site. It is only when you decide to stand up, that Ca\textsuperscript{2+} will be needed and ATP, via the glycolytic pathway and mitochondria, will be delivered. Which brings us to a lot of questions that have yet to be resolved. Let's start from the beginning.

Paper II was primarily a benchmark accessing how the accuracy of the wavenumbers of phosphate molecules', P-O\textsubscript{T} stretching vibrations, depends on different basis sets. The results showed that a well optimized geometry is important and further underlined the importance of the environment being well described. An interesting question is therefore risen. If the wavenumbers of the phosphorylated molecules are better described in a protein environment compared to if the molecules are surrounded by a continuum model. This is a difficult question, since it is computationally very expensive to calculate frequencies on large models with more than a 100 atoms, even using smaller basis sets. Furthermore, to perform such a comparison, the frequencies of several active site models need to be calculated and compared to experimental results where the frequencies are known.

In Paper I, the infrared spectrum of phosphoenol pyruvate (PEP) was studied. The results showed that it should be possible to study the reactivity of PEP in pyruvate kinase using infrared spectroscopy. However, infrared studies did not see such an effect, which could be due to the fact that the reactive conformation is only adapted when ADP binds. The next natural step would therefore be to study the energy profile for the dephosphorylation of PEP in pyruvate-kinase, and to perform frequency calculations on its different conformations.
6. Conclusions and Outlook

In Papers III-V, E2P of the Ca\(^{2+}\)-ATPase was studied through several different perspectives. This approach enabled the deeper understanding of Ca\(^{2+}\)-ATPase’s faster hydrolysis compared to the model molecule, acetyl phosphate in water. Nevertheless, one important key question remains unanswered; is the mystical band at 1115 cm\(^{-1}\) really E1P? The next step would therefore be to study the phosphorylation of the E1P state and to calculate its infrared spectrum.

Hopefully these questions will be resolved in the near future. However, I am sure that just as many new questions will pop up once these are solved.

... For that is the Beauty of the Bitter Devils.
IF you can keep your head when all about you
Are losing theirs and blaming it on you,
If you can trust yourself when all men doubt you,
But make allowance for their doubting too;
If you can wait and not be tired by waiting,
Or being lied about, don't deal in lies,
Or being hated, don't give way to hating,
And yet don't look too good, nor talk too wise:

If you can dream - and not make dreams your master;
If you can think - and not make thoughts your aim;
If you can meet with Triumph and Disaster
And treat those two impostors just the same;
If you can bear to hear the truth you've spoken
Twisted by knaves to make a trap for fools,
Or watch the things you gave your life to, broken,
And stoop and build 'em up with worn-out tools:

If you can make one heap of all your winnings
And risk it on one turn of pitch-and-toss,
And lose, and start again at your beginnings
And never breathe a word about your loss;
If you can force your heart and nerve and sinew
To serve your turn long after they are gone,
And so hold on when there is nothing in you
Except the Will which says to them: 'Hold on!'

If you can talk with crowds and keep your virtue,
'Or walk with Kings - nor lose the common touch,
if neither foes nor loving friends can hurt you,
If all men count with you, but none too much;
If you can fill the unforgiving minute
With sixty seconds' worth of distance run,
Yours is the Earth and everything that's in it,
And - which is more - you'll be a Man, my son!

Rudyard Kipling (1865-1936)

(My grandmother asked me, my siblings and my cousins to learn this poem by heart. She payed us all 50 SEK if we did. I never did.)
7. Acknowledgement

Finally I am sitting down thinking about all the people that have helped me along the way. I am not always the best at expressing my feelings but just thinking about all the wonderful people I would like to mention here, is giving me tears. So to start with I just want to say that I feel incredibly special and fortunate to have you all around me. There are many people that are not mentioned here but that doesn't mean that you are forgotten.

Andreas and Margareta, my two supervisors that gave me the chance to do this work and have given me a lot of freedom, so that my scientific awareness has expanded. Andreas, we have had a lot of wild discussions with a lot diverse point of views, and I really appreciate that you have taken your time to let these discussions have there development. These hours of talking, I have been forced to think my thoughts through properly, therefore it is probably these times that have been the most educational. Margareta, especially during these past months you have been like a second mother to me. You have truly seen me at the worst and let me cry without any judgement, making me feel that you are at my side every step of the way. I have no words to express my gratitude but you should know that it means a lot to me.

During the past years a lot of people have given me inspiration and helped me see the true beauty of quantum mechanics. Starting at Stockholms University, Sten was the person that took me under his wings and helped me every step of the way. His love for science and his pedagogical talents have been a true inspiration. Other big inspirations are Fahmi, Johannes, Holger, Valentin and Katarina. Fahmi, when Sten left for Gothenburg, you were the one I could talk to about my research but also about life. You are a true friend! Johannes, at times when research is all but fun you are the person to turn to – it is amazing what an influential force you have – it only takes you a minute and I am loving it all over again. Holger, you have really proven to
me what a friend is, all your help and effort will never be forgotten! Valentin, first of all, thank you for convincing me to buy a mac, I admit it was the best choice! I really hope we are baking pepparkakor this year too! Katarina, the one who has been on my side through good and bad times. Everybody needs a person on their side that will help them through a PhD and you have really been that person, thank you!

Andreas, Margareta, Johannes, Valentin, Holger, Eeva-Liisa and Emma – thank you all for the time that you spent reading my thesis!!!! Thank you to Maria-Andrea for helping me in Berlin. I would also like to thank Daniel Spångberg for being another big inspiration within quantum chemistry. You gave me the best background in QM one could ask for!

Eeva-Liisa, we have had some great talks both about life at the university and outside university. You have really been the person in our group whom we all could depend on scientifically and non-scientifically. Saroj, dear friend, again thank you for putting up with me when I wanted one more spectrum. Don’t forget that you are a great guy and I know you will do great in the future! Nadja, we had some great times, especially in Croatia – I will never forget the singing boys in high heels. It is great to see your smile – keep it up!

A big thank you to all the past and current members of Andreas B. Group and Per S. & Margareta B. group that I have not mentioned. I would also like to thank the whole of DBB, especially past and current members of the C4-corridor for making life enjoyable at work! It was never difficult to get up in the mornings! Thank you Lena, Haidi and Torbjörn for always helping with a smile.

Before starting my PhD, I didn't know what a hobby was but thanks to a few people, sports is now my biggest hobby. The one person who has stood by my side almost every week these past four years and who probably knows
me better than most, is Mats. Thank you Mats for making me physically strong! Sten, thank you for introducing me to “innebandy” – it is a drug today (and thank you Mats and Johan for making me believe that you want me there). Aron – I could not ask for a better badminton partner! You have also become a true friend! Anders, thank you for showing me that I have a higher limit, than I admit in having. Johanna, a warm thank you for pointing me into the right direction. Half a year ago I was a wreck and today life is so enjoyable! Leelah.

Good friends are people that you know will stand by your side no matter what. Jossan and Henke – I can not ask for better friends, you are the best! Alex, today you are one of my oldest friends and I love hanging out with you and Anna. Bengt och Patrik, if anybody can put a smile on my face any day of the week, it is the two of you! Charlotta (Kära-Lotta), thank you for letting me complain these past weeks or months, it means a lot to me!

Everybody who has a wonderful family is fortunate. Since I have two, it makes me doubly fortunate. Soroush, Kavoos, Jonas and PG – you guys will always be my Uppsala family and I love you all as you were my sister and brothers. Soroush, “livet är bara härligt”, we did it!!! Nobody can ask for a better best friend than you!

And then to my larger amazing family – what would I do without you! Emma, you are not only a sister but a cousin and one of my absolutely best friends. Thank you for showing me that you care. Lasse the best thing you have taught me is “vi börjar stenhårt och ökar successivt”, Sofia, belle-soer and amazing friend. Johannes and Gabriel, you have always been my heroes and it took me years to understand that you weren't perfect! Anna, my dear sister who means the world to me! Mikael, Elisa, Benjamin, Gabriel, Jakob och Rafael – you guys rock my world! Mamma and Pappa, thank you for being AMAZING parents and for helping me no matter what. Andreas, my biggest inspiration of all when it comes to life, brorsan jag älskar dig!
8. Svensk Populärvetenskaplig Sammanfattning


Liksom de flesta enzymer så har kalciumpumpen ett aktivt centrum, det är här som energin i form av ATP utnyttjas. Det som sker är att ett av ATPs tre fosfater frigörs och energi utvinns. Fosfater binds då till ett av enzymets många aminosyror som heter asparaginsyra och bildar aspartylfosfat. En intressant detalj med kalciumpumpen är att det krävs mindre energi för aminosyran att bli av med dess fosfatgrupp i enzymet jämfört med om aspartylfosfatet hade funnits i vatten.


För att muskelceller ska kunna kontraheras så krävs förutom kalcium, ATP. ATP produceras framförallt från cellens energikällor, mitokondrierna. För att mitokondrierna ska kunna bilda energi så krävs både syre och pyruvat.

Brandmannen har ätit upp sin pasta, hans ämnesomsättning eller matspjällkning har satt igång, dvs kolhydraterna finfördelas till glukos. I praktiskt taget alla kroppens celler så bryts glukos ned till pyruvat, och det är just denna molekyl som mitokondrierna kräver. Processen där glukos bildar pyruvat kallas glykolysen och i dess sista steg så bryts phosphoenolpyruvat, PEP, ner till pyruvat i ett enzym som kallas pyruvatkinas. I Artikel I så tolkar jag experimentella resultat på just PEP som sedan kunnat utnyttjats i studier av PEP när den binds till pyruvatkinas.

Fosfat molekylerna har studerats teoretiskt för att kunna förstå molekylernas egenskaper i detalj. Kemiska bindningar bildas när elektroner, en av våra minsta och mest fundamentala enheter, omfördelas. Kvantmekaniken är en
teori som formulerades i början av 1900-talet och som beskriver en partikels position enligt en sannolikhetsfördelning. När man studerar elektroner eller en kemiska bindning så krävs just kvantmekanik. Denna teori är dock endast exakt om man studerar en enstaka partikel och därför krävs approximationer. En sådan approximation som jag utnyttjar i denna avhandling kallas basset. När approximationer utnyttjas är det också viktigt att validera metoden, vilket jag gör i Artikel II.

9. References

(2) Kumar, S.; Barth, A. *Biophys. J.* 2010, 98, 1931-1940.
(14) Barth, A. *Spectroscopy* 2008, 22, 63-82.
9. References


(40) Frisch, M. J. et al Gaussian 09, Revision A.2, Gaussian, Inc., Wallingford CT; 2009.

(41) Jaguar; version 7.6, Schrodinger, LLC, New York, NY; 2009.


9. References

(64) Toyoshima, C. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 2009, 1793, 941-946.
(70) de Marothy, S. A. XYZViewer version 0.963; 2011.