



UNIVERSITA' DELLA CALABRIA

Dipartimento di Chimica e Tecnologie Chimiche - CTC

Scuola di Dottorato

Scienze e Tecniche "Bernardino Telesio"

Indirizzo

Metodologie Chimiche Inorganiche

Con il contributo di

Dipartimento di Fisica

CICLO

XXVII

TITOLO TESI

Study of Enzymatic Mechanisms Using Density Functional Theory and QM/MM Methods

Settore Scientifico Disciplinare CHIM/03 CHIMICA GENERALE E INORGANICA

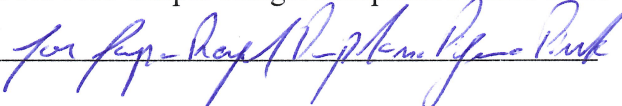
Direttore: Ch.mo Prof. Roberto Bartolino

Firma  _____

Supervisore: Prof.ssa Marirosa Toscano

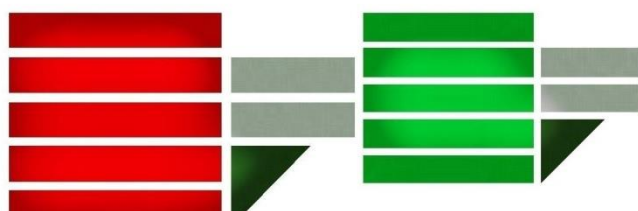
Firma  _____

Dottorando: Dott. Jose Gaspar Rangel Pamplona Pizarro Pinto

Firma  _____

Università degli Studi della Calabria

UNIVERSITÀ DELLA CALABRIA



Dipartimento di Chimica e Tecnologie Chimiche

*Study of Enzymatic Mechanisms Using
Density Functional Theory and
QM/MM Methods*

José Gaspar Rangel Pamplona Pizarro Pinto

11/12/2015

Riassunto

Il lavoro presentato in questa tesi riguarda il meccanismo catalitico di tre enzimi naturali e di un biomimetico. All'inizio del documento è riportata una breve introduzione sul ruolo svolto in natura dalle proteine biologiche in generale, e in particolare dagli enzimi. Poche righe sono state dedicate all'importanza di alcuni metalli di transizione che si trovano più spesso nei siti attivi degli enzimi, con un particolare riferimento ai metalli presenti nei sistemi oggetto di questo studio. Nel capitolo successivo sono stati descritti i principi base delle varie metodologie teoriche utilizzate (QM e QM/MM) ed è stata fatta una breve descrizione della teoria cinetica enzimatica. Data la complessità dei sistemi in studio, dovuto alle loro dimensioni e alla presenza in essi dei metalli di transizione, è stato ritenuto utile dedicare la terza parte di questa tesi ai protocolli computazionali maggiormente usati per trattare argomenti di questo tipo che permettono di ottenere risultati affidabili.

Nei capitoli successivi sono stati riportate le singole problematiche affrontate e i risultati finora ottenuti che hanno riguardato in particolare il meccanismo di lavoro dei seguenti enzimi:

- modello biomimetico del sito attivo dell'enzima 3-Hydroxyanthranilate-3,4-dioxygenase (HAD).
- Tyrosine Ammonia-Lyase (TAL)
- Human Pancreatic Amylase (HPA)
- Organo-Phosphate Degrading enzyme (OPDA)

Abstract

The work presented in this thesis concerns the catalytic mechanism of three natural enzymes and a biomimetic. At the beginning of the document provides a brief introduction on the general role of biological proteins, and in particular on that of enzymes. A few lines are dedicated to the importance of some transition metals that are found more often in enzyme active sites, with a special emphasis to metals present in the systems object of this study. In the next chapter were described the basic principles of the various theoretical methodologies used (QM and QM/MM), and a brief description of enzymatic catalysis theory. Given the complexity of the studied systems, due to their size and the presence in them of the transition metals, it was deemed useful to devote the third part of this thesis to the computational protocols mainly used to treat such arguments that allow us to obtain reliable results.

In the following chapters were given individual issues addressed and the results so far obtained involving in particular the action mechanism of the following enzymes:

-Biomimetic model of the active site of the enzyme 3-Hydroxyanthranilate-4-3 dioxygenase (HAD).

Tyrosine Ammonia-Lyase (TAL)

-Human Pancreatic Amylase (HPA)

-Organo-Phosphate Degrading enzyme (OPDA)

Resumo

O trabalho apresentado nesta tese refere-se ao mecanismo catalítico de três enzimas naturais e um mecanismo biomimético. No início do documento ha uma breve introdução ao papel desempenhado pelas proteínas de natureza biológica, em geral, e em particular pelas enzimas. Algumas linhas foram dedicadas para a importância de alguns metais de transição que são encontrados com mais frequência nos centros ativos de enzimas, com uma referência especial para os metais presentes nos sistemas em estudo. No capítulo seguinte, têm sido descritos, os princípios básicos das diversas metodologias utilizadas teóricos (QM e QM / MM) e também uma breve descrição da teoria de cinética enzimática. Dada a complexidade dos sistemas em estudo, devido ao seu tamanho e à presença neles de metais de transição, foi considerado útil para dedicar a terceira parte desta tese protocolos computacionais mais comumente usadas para tratar este tipo de argumentos que permitem a obtenção de resultados confiáveis .

Os capítulos seguintes são dirigidos aos resultados obtidos até agora. Estes resultados centram-se no mecanismo de funcionamento das seguintes enzimas:

- modelo biomimético do centro ativo da enzima 3-Hydroxyanthranilate-3,4-dioxygenase (HAD).
- Tyrosine Ammonia-Lyase (TAL)
- Human Pancreatic Amylase (HPA)
- Organo-Phosphate Degrading enzyme (OPDA)

Table of contents

Título e composição do Júri	ii
1 Introduction	1
1.1 Enzymatic Catalysis generalities	3
2 Theoretical Background	7
2.1 Quantum mechanics (QM)	7
2.1.1 <i>Density Functional Theory</i>	7
2.1.2 <i>Density Functionals</i>	9
2.2 Basis set	10
2.3 Molecular Mechanics (MM)	11
2.3.1 General Force Field Equation	12
2.3.2 AMBER	14
3 Modeling Reactions	17
3.1 Cluster Model Approach	17
3.2 QM/MM Approach – ONIOM	18
4 A QM/MM study and MD simulation to uncover the reaction mechanism of the Dipeptidyl Peptidase III enzyme	21
4.1 Abstract	21
4.2 Introduction	21
4.3 Methods	24
4.3.1 Models	24
4.3.2 Molecular Dynamics Simulations	24
4.3.3 QM/MM Calculations	25
4.4 Results	26
4.5 Conclusion	33
4.6 Bibliography	34

5	Assessing the promiscuity of the Organo-phosphate degrading enzyme from <i>Agrobacterium radiobacter</i>	37
5.1	Abstract	37
5.2	Introduction.....	38
5.3	Methods	41
5.4	Results	43
5.4.1	Phosphotriesterase mechanism:	43
5.4.2	Promiscuous phosphodiesterase mechanism:	49
5.5	Bibliography.....	57
5.6	Supporting Information	61
6	Scope - Tyrosine Ammonia-Lyase given by a QM/MM model	73
7	Scope – Human pancreatic α -amylase	75
8	Scope - Organo-Phosphate Degrading Enzyme	77

1 Introduction

Proteins can be found virtually in every body part, cells and tissue. They are large and complex molecules and are required for functionality, structure and regulation of the body organs and tissues. They are responsible for most of metabolic processes that sustain life, so the correct work of enzymes is of extreme importance in every human organisms but also in all kinds of organisms, from bacteria to animals.

Enzymes work by lowering the activation energy of a given reaction. The enzyme may do its job by stabilizing the transition state or even by distorting the substrate to the point that the transition state does not require an impossible high Gibbs free energy. Another way in which an enzyme may lower the activation energy of a reaction is by changing its pathway. β -galactosidase deficiency is an example of the importance a single enzyme has in the normal life of a person. Lactose intolerant people do not produce enough β -galactosidases or the ones they do produce work deficiently. This makes lactose impossible to digest at normal enzymatic digestive velocities leaving lactose to be metabolized by bacteria resulting in the fermentation and its product gases.

The active site of an enzyme may be surrounded by just amino acids, or amino acids and co-factors, which help the reaction by stabilizing substrates, transition states or products. In the active site of enzymes, we can also find metal atoms (one or two) or even a cluster of metal atoms. The iron-sulfur cluster enzymes are a good example of the latter.

Enzymes with non-metal active sites are the most common biologically. The substrate enters the pocket where there are electrostatic interactions or covalent bonds fixing-the substrate in the correct position and place so catalysis may happen.

About a third of all proteins are metalloproteins in which the metal atom participates directly or indirectly in the catalysis.

There are ten metals which are commonly found in proteins, either in their cationic form or associated with amino acids. Of the alkali metals, sodium (Na) and potassium (K) and alkaline earth metals, magnesium (Mg) and calcium (Ca) are among the ten

most common metals typically found in the cationic form. The transition metals, manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu) and zinc (Zn), complete the group of metals commonly found in proteins, and are normally associated with amino acids. In addition, manganese, iron and cobalt are also found as components of cofactors such as chlorophyll and heme groups.

Iron is the most common transition metal in biological systems; it is also the most common metal on earth's crust. A complex system, in humans, regulates iron equilibrium. Both excess and scarcity of iron on the human organism has some kind of disease as a consequence. One well known effect of iron deficiency is anemia.

Iron is most known for heme and heme like ligands. But iron-sulfur complexes are also well known and very interesting. Such clusters are found on nitrogenase enzymes with the biological role of nitrogen fixation. Because of a high iron presence on earth it is only natural to try and use iron when synthesizing a biomimetic catalyst. In this document we will show a biomimetic inorganic catalyst, mimicking a ring cleaving enzyme.

Zinc has a stable oxidation state of $2+$ with its d-orbital full. In this oxidation state every electron is paired and as a result it has a low spin configuration. After Iron, Zinc it's the second most common transition metal in proteins. In biological systems its common ligands are mainly histidine residues, aspartic and glutamic acids. Its coordination spheres may go from three to six ligands, though four is the most common number of ligands in a zinc coordination sphere.(1; 2)

Cobalt is not as common in biological systems as both previous referred metals. It may be found in proteins in three different oxidation states, from Co(I) to Co(III). In vitamin B12, a vitamin that plays a key role in controlling the normal functioning of the brain and the nervous system, cobalt has an oxidation state of Co(III). It may also work as a red-ox agent changing from Co(I) to Co(II) and vice versa. It is also found in binuclear centers as we'll see farther in this document.

1.1 Enzymatic Catalysis generalities

A catalyst is a substance that increases the rate of a reaction without itself being consumed by the process. It lowers the Gibbs energy of activation by providing a different mechanism for the reaction. This mechanism enhances the rate and it applies to both the forward and the reverse directions of the reaction.(3)

The catalyst forms an intermediate with the reactant(s) in the initial step of the mechanism and is released in the product-forming step, thus does not appear in the overall reaction. Regardless of the mechanism and the energetics of a reaction, a catalyst cannot affect the enthalpies or Gibbs energies of the reactants and products. Thus, catalysts increase the rate of approach to equilibrium, but cannot alter the thermodynamic equilibrium constant.

There are three types of catalysis: heterogeneous, homogeneous, and enzymatic. Enzyme catalysis is also mostly homogeneous in nature. However, because it is of biological origin is the most complex of the three types of catalysis, and it is treated as a separate category.

An enzyme usually contains one or more active sites, where reactions with substrates take place. An active site may comprise only a few amino acid residues; the rest of the protein is required for maintaining the three-dimensional integrity of the network. The specificity of enzymes for substrates varies from molecule to molecule. Many enzymes exhibit stereochemical specificity in that they catalyze the reactions of one conformation but not the other. Some enzymes are catalytically inactive in the absence of certain metal ions. Although enzymes are evolved to catalyse a particular reaction on a particular substrate with a high catalytic efficiency (k_{cat}/K_M , Michaelis–Menten kinetics), they can often perform side reactions in addition to their main, native catalytic activity. These promiscuous activities are usually slow relative to the main activity and are under neutral selection.(4) Despite ordinarily being physiologically irrelevant, under new selective pressures these activities may confer a fitness benefit therefore prompting the evolution of the formerly promiscuous activity to become the new main activity. Promiscuity comes in many different forms. We can have the so called “enzyme condition promiscuity” that is shown by enzymes with catalytic

activity in various reaction conditions different from their natural ones, such as anhydrous media, extreme temperature or pH and so on. Another type of promiscuity is known as “enzyme substrate promiscuity” and it is shown by enzymes with relaxed or broad substrate specificity. Finally, we have the “enzyme catalytic promiscuity” which is shown by enzymes catalyzing distinctly different chemical transformations with different transition states. The latter can be either:

- (i) accidental – a side reaction catalyzed by the wild-type enzyme;
- (ii) induced – a new reaction established by one or several mutations rerouting the reaction catalyzed by the wild-type enzyme.

In the 1890s the German chemist Emil Fischer (1852–1919) proposed a lock-and-key theory of enzyme specificity. According to Fischer, the active site can be assumed to have a rigid structure, similar to a lock. A substrate molecule then has a complementary structure and functions as a key. Although appealing in some respects, this theory has been modified to take into account the flexibility of proteins in solution. We now know that the binding of the substrate to the enzyme results in a distortion of the substrate into the conformation of the transition state. At the same time, the enzyme itself also undergoes a change in conformation to fit the substrate.

The enzyme and substrate interact to form an enzyme-substrate complex. The interactions between the substrate and active site are weak, noncovalent interactions (i.e. the substrate does not covalently bind to the active site but weakly interacts with it through interactions like hydrogen-bonding, van der Waals interactions, etc).⁽⁵⁾⁽⁶⁾ The orientation in which the two interact is highly favorable for facilitating conversion of the substrate to product. In the enzyme-substrate complex, the substrate molecule binds to a very specific region of the enzyme molecule called the active site. The active site is usually found in a 3-D groove or pocket of the enzyme, lined with amino acid residues (or nucleotides in RNA enzymes). These residues are involved in recognition of the substrate. Residues that directly participate in the catalytic reaction mechanism are called active site residues. After an active site has been involved in a reaction, it can be used again. Substrates bind to the active site of the enzyme through hydrogen bonds, hydrophobic interactions, temporary covalent interactions (van der Waals) or a

combination of all of these to form the enzyme-substrate complex. Residues of the active site will act as donors or acceptors of protons or other groups on the substrate to facilitate the reaction. In other words, the active site modifies the reaction mechanism in order to change the activation energy of the reaction. An enzyme binding to a substrate will lower the energy barrier that normally stops the reaction from happening. The product is usually unstable in the active site due to steric hindrances that force it to be released and return the enzyme to its initial unbound state.

As we will see, the active site in an enzyme is of crucial importance to study the catalytic behavior from a computational point of view.

2 Theoretical Background

2.1 Quantum mechanics (QM)

The first approximation, applied in any theoretical method, is certainly the Born-Oppenheimer approximation, according to which, given the difference in mass between the nuclei and electrons, it is possible to decouple the motion of the nuclei from that of the electrons whilst the nuclei are considered as fixed. By breaking down the molecular system into two subsystems (electronic and nuclear), the expression of the Schrödinger equation is simplified, as it is possible to neglect the kinetic energy of the nuclei but still maintain the nucleus-nucleus repulsive interaction. The search for theoretical methods apt to provide approximate solutions of Schrödinger equation without resorting to parameters obtained experimentally has led to the development of Hartree-Fock (HF), post-Hartree-Fock (post-HF) and Density Functional Theory (DFT) methods. With the methods previously mentioned we are able to run computational studies of a few hundred atoms. Despite the higher theoretical level of the HF and post-HF methods, in this thesis we use and present only the DFT method for its lower computational costs while maintaining a high theoretical level and accuracy. The density functional theory states it is possible to extract the properties of a given system from its electronic density. (7)

2.1.1 Density Functional Theory

The density functional theory foundation lies in the two Hohenberg-Kohn theorems(8). Hohenberg and Kohn showed that the density of the ground state of a non-degenerate electron system makes it possible to uniquely determine all the properties of the ground state. So each physically observable ground state, such as the total energy of the ground state, can be expressed as a functional of the density, $E[\rho]$. The energy $E[\rho]$ is in fact a function of the electron density $\rho(\mathbf{r})$, which is in turn a function of the spatial coordinates (x, y, z) . The electronic energy functional $E[\rho]$ is generally divided into several contributions which take into account separately the electronic kinetic energy T

$[\rho]$, the potential energy of electron-electron repulsion interaction $V_{ee} [\rho]$ and the potential energy of electron-nuclei interaction $V_{Ne} [\rho]$. The first two contributions are independent of the nuclear positions and can therefore be grouped together in the universal density functional $F [\rho]$:

$$E [\rho] = F [\rho] + V_{Ne} [\rho] \quad (6)$$

The exact determination of the functional $F[\rho]$ is in fact not possible. Because of the limitations of this equation Kohn and Sham introduced a set of orbitals (analogous to HF orbitals) in which the kinetic energy can be divided into two parts. [9]

In their theorem, Kohn and Sham, represented the functional of the kinetic energy of a system in two parts, one which could be exactly calculated assuming that the electrons do not interact with each other, and another part which was a correction to account for the interaction between them.

The Kohn-Sham formalism states that the ground state electronic energy in a system, with n electrons and Z nuclei is calculated as follows:

$$E[\rho] = -\frac{1}{2} \sum_{i=1}^n \int \Psi_i^*(r_1) \nabla_i^2 \Psi_i(r_1) dr_1 - \sum_{X=1}^Z \int \frac{Z_X}{r_{Xi}} \rho(r_1) dr_1 + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + E^{XC}[\rho] \quad (7)$$

The four terms in the equation are respectively kinetic energy of electrons where there are interactions, Ψ represents the orbital Kohn-Sham, the attraction between electron density and nuclei, the Coulomb repulsion distribution of the total load r_1 and r_2 and the term exchange - correlation which represents the correction made to the first term where there are no interactions between the nuclei. Here the exchange functional describes the exchange interaction between the electrons with the same spin

multiplicity and functional correlation describes the interaction of electrons with opposite spin multiplicity.(9; 10)

2.1.2 Density Functionals

The density functionals first developed were very simplistic and two different types of density functional existed at the time. The big difference between them was the inclusion of spin effects in one type (LSDA – Local Spin-Density Approximation)(11; 12; 13; 14) and the other type of density functionals had no inclusion of spin effects (LDA – Local Density Approximation)(15). Density functionals such as the GGA (Generalized Gradient Approximation) (e.g. BLYP and PBE)(16; 17) that perform better than the earlier LDA and LSDA are now in use. There are also some further developments of this type of orbital, when calculating the second derivative of electron density, which is the kinetic energy density, the functional Meta - GGA (e.g. TPSS and M06-L)(18) results can be obtained with higher accuracy. There are also functional hybrids can be H-GGA (e.g. B3LYP) or HM-GGA (e.g. TPSSh)(19), based on whether or GGA to M-GGA respectively(20; 21; 22; 23).

The development of density functionals is still ongoing with new functionals every year.

2.1.2.1 B3LYP

B3LYP is a hybrid density functional that was obtained from the combination of other the exchange and correlation functionals. In addition to the introduction of B88, LYP and VWN it was also included 20% of exact HF exchange(22; 24).

$$E_{XC}^{B3LYP} = (1 - a)E_X^{LSDA} + aE_X^{HF} + bE_X^{B88} + cE_C^{LYP} + (1 - c)E_C^{VWN} \quad (8)$$

In equation 8 the coefficients $a=0.20$, $b=0.72$ and $c=0.81$ were obtained from the B3PW91 hybrid density functional.

In this work we chose B3LYP to all optimizations given that it has been broadly tested and used and is considered to yield good geometries. The parameters introduced in this functional were fitted to energies of atoms from the three first rows of the periodic table, however, studies have shown that geometries of various transition metals very small mean absolute deviation.

2.2 Basis set

In the case of the method that was briefly described above as well as in other *ab-initio* and DFT methods, basis-set functions are needed for this calculation. Basis-set functions allow us to describe approximately an unknown function.

Including basis-set is not in itself an approximation, assuming that the unknown function is fully described, but for that to be true an infinite number of basis-set, which is an impossibility in practical terms. Obviously the smaller the number of basis-set functions, the worse will be the representation of an unknown function. Other than the number of basis-set used to describe an unknown function the correct choice of basis-set used is also of the utmost importance.

There are two common types of basis-set functions used: STO - Slater Type Orbitals and GTO - Gaussian Type Orbitals.

The Slater-type orbitals describe more accurately the electron- nucleus interaction.(25; 26) However, for non-hydrogenoid molecules, this orbital is not suitable as it computational expensive to use it for large molecules.

The Gaussian type orbitals have a simpler mathematical resolution. However their description of the behavior of electrons is not done as well, especially when the electrons are near the core. Nonetheless, for large molecules their use is considered to be the better choice.

Whatever the chosen basis-set is, they shall not describe with accuracy only one atom, but be flexible enough so that they can describe the molecule as a whole and not a particular atom. This good description on the molecular level can be obtained with the inclusion of polarization and diffuse functions. The polarization functions have a

higher angular momentum, thus describing the distortion of atomic orbital in a molecular environment. On charged or highly polarizable molecules is advisable to use diffuse functions because they describe best electron density away from the core.

2.2.1.1 Pseudopotentials

When studying systems of considerable dimensions or even transition metals it is wise to use basis-set with pseudopotentials. The introduction of pseudopotentials has the purpose to lower the computational cost of the calculations by lessening the number of functions needed to describe all the electrons present in a calculation.(27)

This methodology is based on the assumption that core electrons and other lower energetic level electrons are mainly inactive and do not disturb valence electrons that are responsible for bond properties.

2.3 Molecular Mechanics (MM)

Molecular mechanics is one of the most used methods in the study of biochemical systems but as for other methods there are advantages and disadvantages. Molecular mechanic calculations are much faster than the quantum mechanical calculations and are therefore used in systems with many atoms. However their major drawback is to be unable to explicitly describe electrons. This method can't be applied directly to study a system in which electrons are exchange between atoms or even a study of how a particular chemical connection is made. The methods used in the calculation of molecular mechanics (MM) are quite different from those used in quantum mechanics (QM). Unlike QM methods that are based in the Schrödinger equation, MM methods describe the energy of a system just from the coordinates of its nuclei. As the MM can't explicitly describe electrons, it considers atoms to be balls connected to each other through springs, where the balls have different sizes and springs have different loads depending on the atom they are representing. Here the energy of the system is a sum of several different processes such as stretching of bonds and bending of angles.

MM methods also require parameters that are obtained from smaller model calculations using a much higher theoretical level such as Hartree-Fock methods. Alternatively, experimental data can also be used. The accuracy of these parameters is therefore of utmost importance to have a high degree of confidence in molecular mechanics studies.

2.3.1 General Force Field Equation

A force field is a set of parameters used to define the potential energy of a system. The set of parameters of a force field comes from experimental work and calculations at a quantum mechanics level. The energy of a force field is usually described as the sum of several energy terms with different physical meaning and contribution. An example of the general equation is as follows:

$$ET = Estr + Ebend + Etors + EVDW + Eele \quad (9)$$

$$Estr = \frac{1}{2}k_l(l - l_0)^2 \quad (10)$$

$$Ebend = \frac{1}{2}k_\theta(\theta - \theta_0)^2 \quad (11)$$

$$Etors = \sum_j \frac{1}{2}V_j [1 - \cos(j\omega)] \quad (12)$$

$$E_{VDW} = \sum_{ij} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right] \quad (13)$$

$$E_{ele} = \sum_{ij} \frac{q_i q_j}{\epsilon r_{ij}} \quad (14)$$

In the equation (9) the first term corresponds to the stretching of bonds, the second term to the bending of angles, the third term describes the torsion of dihedral angles. These are the three terms with bonding interactions. The two following terms describe the non-bonding interactions and correspond to the energy of van der Waals and

electrostatic energy respectively. There are some force fields that include a correlation term.

In equation (10) k_l is the force constant and l_0 is the constant that describes the reference distance between the atoms considered. The approximation resulting from the use of Hooke's formula is usually accepted for bond length variations near the reference value associated with the minimum energy. Even though the Morse potential gives a better description of the energy of a given bond, the harmonic potential yields similar results in the bond range that chemists normally work and it is computationally cheaper. The difference is in the extremes, where the Morse potential describes the bond energy better than the harmonic potential.

As the stretching of the binding energy, the energy of an angle can be described in a harmonic way. In equation (11) k_θ is the force constant of the angle and θ_0 the reference value of the angle at the balance position.

The torsion energy is the energy required for the rotation of the bond AB and CD in a group of atoms designated ABCD. Unfortunately, unlike the other two terms of binding interactions, this term is more difficult described by a harmonic potential. This is due to the fact that torsion energy presents a cyclic behavior and the variation between each minima requires very low energy. In equation (12), ω is the dihedral angle, i. e., the angle that AB makes with CD, j is its multiplicity, the minimum number of function while the link is rotated 360° (e.g. if $j = 6$, the function will have five minima and each energetic minima will be separated by 60°) and V_j is the torsional constant force.

The van der Waals term describe interactions between atoms which are not bonded, regardless of their charge. These interactions are attractive for small distances but tend to null very quickly when the distance between atoms increases. At small distances, i.e., when the electronic clouds overlap there is van der Waals repulsion. This interaction is commonly calculated using the Lennard - Jones potential. When seeing equation (13) we may be lead on to think that the use of 12 and 6 on the repulsion – attraction potencies is the one that describes best the van der Waals term. It is just

another approximation to lower the computational cost. In the equation 8, A and B are constants, and r_{ij} is the distance between the two atoms.

The electrostatic term in the total energy equation, equation (14), describes the non-bonding interactions between positively and negatively charged atoms; usually this interaction is described by the Coulomb equation. Electrostatic energy is a function of atomic charges (q_i, q_j), the distance between these (r_{ij}) and dielectric constant (ϵ).

These terms may differ slightly from force field to force field.

2.3.2 AMBER

The AMBER force field(28; 29; 30; 31; 32) (Assisted Model Building and Energy Refinement) was initially thought to be used in proteins and nucleic acids. Nowadays the AMBER family has a variety of parameterization schemes that describe proteins and nucleic acids (FF## libraries). There are also parameters to small organic molecules, especially within GAFF(33; 34) (General Amber Force Field) and also carbohydrates parameters in GLYCAM##.(35)

The AMBER energy calculation is described by this mathematical equation:

$$U(\vec{R}) = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_n}{2} (1 + \cos(n\phi - \gamma)) + \sum_{nonbonding} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right] \quad (15)$$

Where k_b , k_θ and V_n are the force constants, b , θ and ϕ are the bond distance, angle and dihedral angle, with subscript zero indicating the reference values for these quantities. In this equation γ is the phase angle can have values between 0 and 180. In non-bonding terms A_{ij} B_{ij} are the van der Waals dispersion, q_i q_j are the partial charges of atoms and ϵ_i is the dielectric constant. The nonbonding interactions terms of the equation are calculated for atoms separated by three or more atoms.

From the computational point of view, metalloproteins are a very particular case. The presence of a metal atom can cause problems in applying some methods that are

already well established and proven to give good results in proteins without metal active sites. These problems can be very easily seen when needing to perform a molecular dynamics simulation on a metalloprotein, where the lack of specific parameters makes their computational study various steps behind the computational study of other proteins.

This is therefore a major limitation of the force fields and also obviously of the AMBER force field.

3 Modeling Reactions

Given that there are two different set of physical laws for the computational study of catalysis mechanisms there are also two different approaches for doing it. Depending on what one is studying it is better to use on approach or the other. There is not one better than the other though there is those who defend one over the other. For instance, because DFT methods are based on quantum mechanics laws, the electrons are taken into account, but the computation of these methods take a very long time, and the size of the model used does not usually exceeds two hundred atoms. On the other hand the MM methods are based on classical mechanics laws, which do not take electrons into account. The smallest unity in MM methods is the atom and it is connected to other atoms with a spring coil. Though they cannot describe chemical reactions they are much faster than DFT calculations and as consequence the size of the model used may be in the thousands. When used alone both approaches have different goals. Since classical mechanics do not take electrons into account, classical mechanic laws cannot be used to describe a chemical reaction with atomistic detail. Though they can and are used in molecular dynamics which gives us an insight into how a given protein moves in time. With quantum methods alone and today's technology it would be impossible to study the dynamics of a protein.

So given the limitations of each approach we know that we have to do some compromises.

3.1 Cluster Model Approach

In the cluster model approach we use *ab-initio* methods (DFT in our case). As we said before we cannot use many atoms and it follows that the whole protein is out of question. The compromise here is using a small part of a given protein to study its catalytic mechanism. This small model of the enzyme will be of the active site of the enzyme, where the catalysis occurs.

When we have a crystallographic file of the whole enzyme, the active site has to be identified as well as the most important amino acids, those that have an important role in the catalytic mechanism. Ideally the model will have the amino acids that participate

in the catalytic mechanism and some others that stabilize the ones that do participate in the reaction. Some may ask if the rest of the enzyme is not important to the mechanism. The idea behind the cluster model approach is that the chemical reaction is concentrated in the active site and that the long range interactions are too small to have such a big contribution for the overall energy.

Nevertheless, after the geometry optimizations for all the stationary states along the reaction path are finished and the enzymatic mechanism is found, single point calculations are done with a polarizable continuum model to simulate the rest of the enzyme or the presence of a solvent. The compromise here is having the same dielectric constant embracing the cluster model instead of having different charges depending on the amino acid present in a given spot of the enzyme, thus eliminating any possible long-range interactions that could stabilize or destabilize the active site.

Also in a cluster model approach the amino acids present in the model should have an atom frozen in space so as not to assume conformations it would not assume with the constraints imposed by the rest of the protein. Normally the atom that stands the farthest away from the active site should be frozen in order to allow as much freedom as possible to the amino acid.

When building a model one must take very careful into account the ratio of the model size per computation time. The model must be consistent enough that the mechanism is well described but also small enough so the computational cost is not exaggerated.

3.2 QM/MM Approach – ONIOM

Other than use only quantum mechanics with a small part of the enzyme (even if it is the most important part for the catalytic mechanism) we can use classical mechanics. For the study of catalytic mechanisms it is possible to combine computationally, quantum mechanics and classical mechanics. With this approach we use *ab-initio* methods in the active site and at the same time maintain the rest of the enzyme to be treated with molecular mechanics. In doing this the active site has to be identified and the amino acids that participate in the reaction mechanism have to be known so they

are included in the QM part of the calculation. In an ONIOM(36) calculation the QM part is usually smaller than it would be if we were doing a cluster model calculation. With ONIOM it is not needed to maintain the supporting amino acids in the QM part of the calculation for they will be present just the same but will be treated with MM. The interface of the layers represents a problem. It would not do to sever the bond and treat one part with QM and the other with MM and after the optimization of the geometry sum the energies.

What happens at the interface of both treatments is that we have to say to the program that the atom that is to be treated at a QM level of theory is bonded to a hydrogen atom. This way there are no atoms with a deficit of bonds. To facilitate the explanation the active site which is treated with *ab-initio* methods will be called the “small model” and the whole system will be called “real system”.

What an ONIOM calculation does is actually divide the job into three steps. On the first step an ONIOM calculation treats the real system with a low level of theory, meaning it uses the parameters in the AMBER libraries to calculate the energy of the enzyme's conformation. After obtaining the energy for the real system, the energy of the small model is calculated, still using a low level of theory. The third step is the calculation of the energy in the small model with a high level of theory, normally and in our case, DFT level of theory is used.

In the end the energy in an ONIOM calculation is obtained as follows:

$$E^{\text{ONIOM}} = E^{\text{low}}(\text{RS}) - E^{\text{low}}(\text{SM}) + E^{\text{high}}(\text{SM}) \quad (16)$$

In equation 16, RS is the real system and SM is the small model.

4 A QM/MM study and MD simulation to uncover the reaction mechanism of the Dipeptidyl Peptidase III enzyme

4.1 Abstract

In this work, the catalytic mechanism of dipeptidyl peptidase III (DPPIII) was studied with atomistic detail, using a hybrid quantum mechanical/molecular mechanical method at the ONIOM (B3LYP/6-31G(d):Amber) level. The hydrolytic reaction proceeds via a general acid/base mechanism, in which the first mechanistic step involves a proton transfer from the zinc-bound water molecule to the Glu508. The second step involves the remaining hydroxyl zinc-bound group that performs a nucleophilic attack on the scissile carbonyl bond leaving the zinc connection to the substrate oxygen atom. The third step involves a proton transfer from Glu508 to the peptide nitrogen and a subsequent cleavage of the peptide bond to yield the products in their neutral forms. The conserved residue Glu508 is ideally aligned and has the ability to slightly rearrange its conformation to act as a highly effective proton shuttle. Our results so far indicate that the nucleophilic attack agrees with the data available for other enzymes with reaction mechanism similar to the DPPIII, with activation energy for the second step being close to 16 kcal/mol.

4.2 Introduction

Zinc containing metallopeptidases constitute an expanding list of structurally related proteases which are widely distributed in nature. They are involved in highly important biochemical events, such as deformylation in bacterial protein synthesis, the case of peptide deformylases, extracellular matrix degradation and tissue remodelling (matrix metalloproteinases), digestion (carboxypeptidase A, astacin), enzyme hypertension management (human angiotensin-converting) and blood-pressure regulation (neprilysin), to name a few examples. Within the big family of enzymes that contain a Zn^{2+} in its active site there is a branch with a special HEXXH binding motif, where the XX stands for different residues in different enzymes. For this reason, different enzyme active sites may even be superimposed and matched based on these similarities.¹⁻⁶ The similarity among the metallopeptidases is such that a standard orientation has been proposed for the visualization of the active site. Since the active site cleft is large enough for a polypeptide to enter, it is proposed that the substrates of these enzymes' family align somewhat horizontally in the cleft. They bound in the active site N to C-terminal.⁷

The HEXXH (or HXXEH and other small deviations) is directly involved in the

catalysis capabilities of the enzymes with the two histidine residues being in the first sphere of coordination and the glutamate residue participating in the mechanism as a general acid/base. Bonded to the zinc atom is also another glutamate residue (E508).^{8,9} The enzyme of our study, dipeptidyl peptidase III (DPP III), was first discovered in the bovine anterior pituitary gland. It has in the active site a zinc atom bonded to two histidine residues (from the motif HELLGH), a glutamate residue and a water molecule. Stabilizing the water molecule is a glutamate residue (from the HELLGH motif).

Later DPP III was shown to be present in several animal tissues, such as rat and human. Both the rat and the human DPP III were proven to have a Zn^{2+} ion per mol of protein. In 2008 the yeast DPP III structure was resolved, showing a different protein fold. Nevertheless the zinc coordination three-dimensional structure and full HELLGH motif is the same as seen in the human and rat DPP III.

This motif is so important that promotes metal promiscuity to the enzyme. A previous mutagenic study as shown that substituting the Zn^{2+} for other metal atoms such as copper (Cu^{2+}), cobalt (Co^{2+}) and nickel (Ni^{2+}) maintains the ability of the enzyme to hydrolyse the substrate Arg-Arg-NA (two arginine residues and a naphthalene group bonded). On the other hand when a leucine residue (L453) is deleted from the enzyme the catalytic ability of the enzyme disappears and the Arg-Na peptide bond is not hydrolysed.^{10,11}

DPP III is an important enkephalin-degrading enzyme associated with the human pain modulatory system. Inhibition studies show that a neuropeptide named spinorphin formed by amino acids (Leu-Val-Val-Tyr-ProTrp-Thr) is effective in the inhibition of monkey purified DPP III.¹² This neuropeptide acts differently than morphine and is widely used as an analgesic in morphine-resistant cases.¹³

In recent years the study on DPP III has grown, with overwhelming results. In the last 6 years structural data on inhibitors and substrate are available, kinetic studies were done to yeast,^{14,15} rat¹⁰ and human¹⁶ DPP. Because of the works and results in literature we now have more data to compare to computational studies.

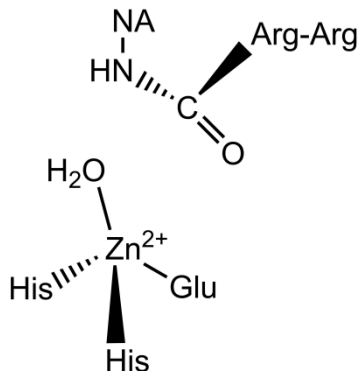
Computational studies have also been done on the DPP III and other enzymes with similar proposed mechanism. As said before the active site cleft is large enough for a polypeptide to enter and accommodate, this means that DPP III hydrolyses a variety of peptides. The conformation that the enzyme acquires when a substrate is bonded to the active site is then, very important for the activity of the enzyme. This data is already in the literature which helped us with our study, as we will show in the results.¹⁷

Also of importance to our work, was a previous QM/MM study on ACE (angiotensin-converting enzyme). This work shows a very similar active site of different enzyme, with a reaction mechanism similar to what we expected to obtain.

Thermolysin, another zinc containing enzyme, was also studied computationally at a DFT level of theory, showing a reaction mechanism very close to the reaction mechanism shown on the ACE work. Moreover, the reaction mechanisms of both the ACE (QM/MM) study and the thermolysin (Cluster Model) study yield very similar energetic results for similar mechanistic steps.

These two reaction mechanisms initiate with a nucleophilic attack of a water molecule on the carbon of the peptide bond (scissile bond). In this process the water molecule

loses a proton that will eventually bond to the nitrogen to break the C – N bond. In both cases the first sphere of the zinc coordination is exactly the same we found in our enzyme, i.e., two histidine residues, a glutamate and a water molecule. These ligands adopt a tilted trigonal pyramidal geometry (scheme 1). The substrate is on top of this pyramid with a residue (histidine or tyrosine) stabilizing the oxygen of the carbonyl group, O_{pep} . Stabilizing the water molecule is a glutamate residue from the HELLGH motif, confirming its importance on the overall catalytic capability of the enzyme.



Scheme 1- Zinc ligands adopted geometry and relative position of the substrate.

4.3 Methods

4.3.1 Models

The DPPIII:substrate complex was modeled from two different X-ray crystallographic structures that were obtained from the PDB databank¹⁸ (ID: 3FVY¹³ and 3T6B with a resolution of 1.90 and 2.00 Å respectively). We first had to superimpose both structures by their catalytic residues because the 3FVY has a zinc atom in active site and the 3T6B does not. On the other hand, the 3T6B structure has an inhibitor from which we modelled the substrate Arg-Arg-Na.¹⁴ After the substrate was modelled, we proceeded to delete the 3T6B structure. The resulting model was used as a starting point for the computational studies. Molecular dynamics (MD) simulations were performed to evaluate the effectiveness of modelling and to confirm that it was stable near the active site. Parametrization was needed for the substrate. The zinc ligands were not needed to be parametrized since that work had been done before by Brás et al.³

For the MD simulations we used the amber force field (ff99sb)¹⁹ parameters for the protein. Explicit solvent TIP3P water was used in a cubic box with a minimum of 12 Å around the protein, and a total of nineteen (19) sodium (Na⁺) ions were added to neutralize the charge of the system. The resulting models consisted of X atoms.

4.3.2 Molecular Dynamics Simulations

We ran three minimization steps, the first one with the protein fixed and the water molecules free to move, in the second step we kept the water molecules free of constraints and freed also the substrate, since it was modelled it had to minimize its energy before the whole system was free of constraints. Finally we proceeded to the third step in which the whole system had no constraints. The minimized structures were then submitted to a further 200 ps long warm-up simulations, from 0 till 300K, at constant volume and using periodic boundary conditions. Further 200 ns Langevin dynamics with 2 fs integration step were carried out. Non-bonded interactions were treated with the PME algorithm and the cut-off was set to 10 Å, real part also truncated at 10 Å with pressure of one atmosphere at 310 K with ensemble NTP. All calculations were performed using the Amber12²⁰ simulations package. Bonds involving hydrogen atoms were constrained with the SHAKE algorithm²¹.

Since the conformations obtained from the MD simulation were in agreement with the conformations said to be most apt for the mechanism to occur¹⁷ we continue our study with the next step.

4.3.3 QM/MM Calculations

We performed QM/MM calculations to determine the potential energy surface (PES), using the Gaussian 09 software²². We took from the last minimization the model used for the calculations. We froze the residues that were farther than 15 Å from the active site and most of the water molecules from the solvent were cut off yielding a model with a total of 11914 atoms and 898 residues. We divided our system in two layers and the ONIOM formalism²³ was used in order to calculate the corresponding PES. In the high layer there are 100 atoms, *i.e.* one zinc atom, the naphthalene group from the substrate, the backbone of the arginine residues of the substrate and two CH₂ of each side chain, one water molecule, two histidine residues side chains and two glutamate residues side chain (until the beta carbon). The high layer was treated with density function theory (DFT) at the B3LYP/6-31G (d) level²⁴⁻²⁷. The low layer (the rest of the enzyme together with the solvent water molecules) was treated at the molecular mechanics level with the parameters of the amber force field package. We used electrostatic embedding to treat the coupling between both layers.

Transition states were searched for using flexible scans. Scans were made along the reaction bond in each step of the mechanism. After knowing the approximated geometry of each transition state, an optimization was carried out to find the geometry of the transition state and respective frequencies. To ensure that the minima found corresponded to same reaction coordinate as the transition state, IRC calculations were made.

The zero point energies were calculated at the same QM/MM level as the geometry optimizations, within the harmonic approximation. Thermal corrections (at 310 K) and entropy were calculated within the ideal gas/rigid rotor/harmonic oscillator model.

Single point energy calculations were performed subsequently with the B3LYP density functional^{28,29} with the larger 6-311++G(2d,2p) basis set.

4.4 Results

Given the importance of DPP III in the pain modulatory system it is underwhelming the amount of data available in the literature. There can be found experimental studies on it. Such as kinetic studies on the bacteria, rat and human DPP III, nevertheless computational studies on the mechanism itself, be them at DFT level or at QM/MM level, are not found, to the best of our knowledge. In the PDB databank for instance, we can only find four DPP III crystallographic structures. Three originated from human and one from *Saccharomyces cerevisiae*. Which is very few compared to the hundred and twenty eight (128) crystallographic structures of thermolysin present in the same databank.

Because of the quantity of Zinc-metalloproteinases and their importance, it is expectable that some of them are not as studied as others at a given time. The example of thermolysin is apparent, but also β -lactamases and angiotensin converting enzymes have been studied and their catalytic mechanisms have been proposed and supported by computational studies.

This study as not yet reached its final stage. The last mechanistic step has not yet been uncovered. However, the first and second steps have been uncovered and the rate limiting step energetics is in agreement with the data obtained experimentally.

Because of the discussion and conclusions about the conformation of DPP III we decided to run a rather long molecular dynamics simulation. In figure 2 we show the RMSd of our MD of 150 nanoseconds, though the RMSd is rather high once it is stabilized, it is explainable through the high mobility of both arginine residues in the substrate. The side chain of arginine is rather long and has six single bonds in a row, meaning it is very movable. The overall conformation and position of the substrate in the active site is stable throughout the MD simulation.

On the other hand the RMSd of the protein's backbone stabilizes around 2.2 angstroms and stays that way throughout the simulation. Nevertheless we decided to see where the higher fluctuations on the protein were. Unsurprisingly we learned that the outer shell of the protein had the highest fluctuations. The active site and the substrate have very low fluctuations (figure 3).

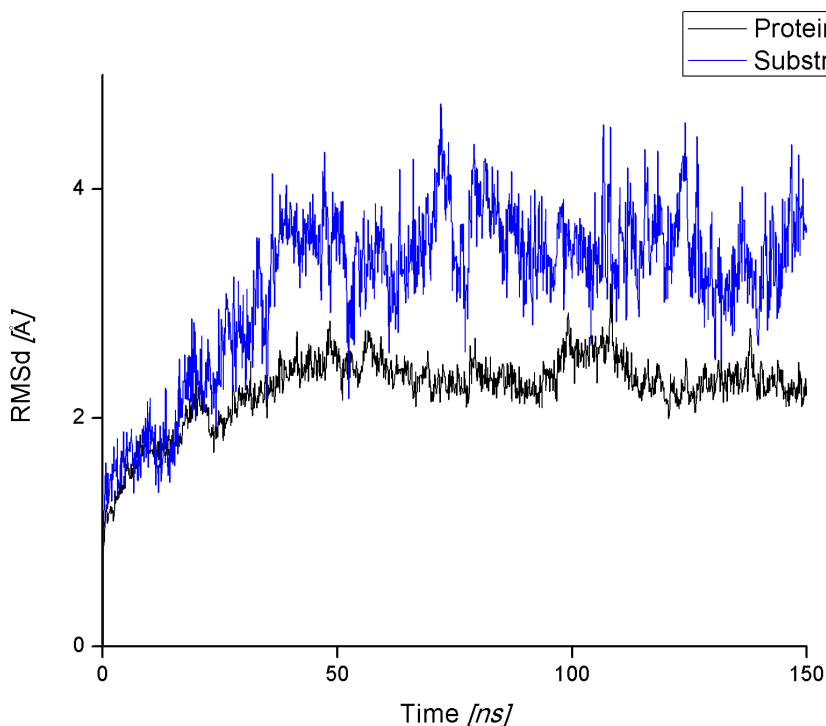


Figure 2- RMSd values for the protein backbone and the substrate.

The reactants model we used to start the calculations (scheme 1) has the Zinc bonded to four ligands. Adopting a trigonal pyramidal conformation are two histidine residues, a glutamate residue and a water molecule. The water molecule forms a hydrogen bond with a glutamate residue. Directly above the water molecule is the scissile carbon of the substrate. Close to the substrate and stabilizing it in the active site is a histidine residue.

From this initial model we knew from similar enzymes that the oxygen of the water molecule would attack the scissile carbon. After trying to scan the attack directly with the water molecule we saw that it was not strong enough to attack and that every time, one proton from the water molecule would dissociate from the water to bond to the glutamate residue (figure 1).

This first step is energetically very cheap, with the transition state and the successive intermediate having virtually the same energy of the reactants. The energy of the three

states is so close to one another because there is not much change in the active site. In the reactants the proton is at 1.437 Å from the glutamate residue and still bonded (1.078 Å) to the nucleophilic oxygen. The transition state shows the proton almost midway between the Glu508 residue (1.211 Å) and the nucleophilic oxygen (1.234 Å). At the end of this mechanistic step, the first intermediate is achieved and the proton is now bonded to the Glu508 residue, with a bond length of 1.075 Å and is forming a hydrogen bond with the nucleophilic oxygen, at a distance of 1.460 Å. Because the nucleophilic oxygen is bonded to the zinc the bond it formed with the proton in the reactants conformation was already elongated for a normal O – H bond in a water molecule (0.998 Å) in liquid phase.³⁰ This elongation of the bond, allows the transition of the proton from one oxygen atom to the other to have no energetic expense (figure 4).

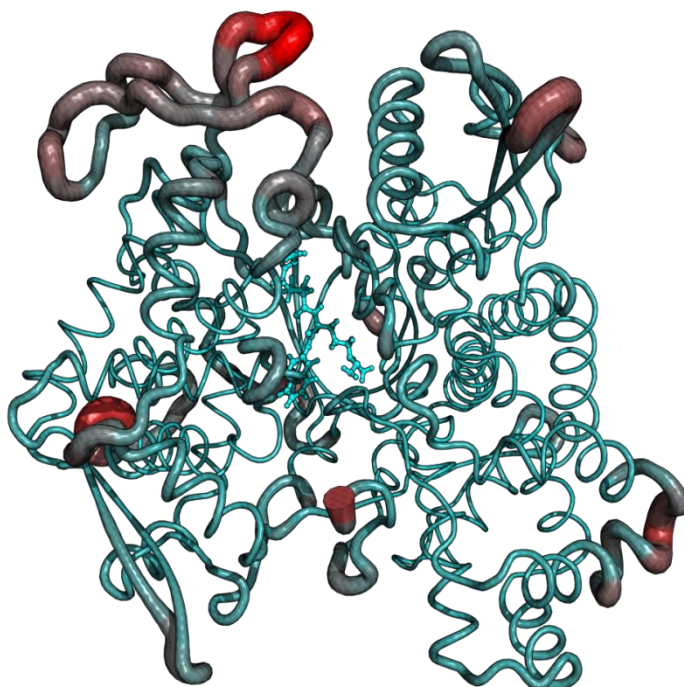


Figure 3- Root mean square fluctuations of the enzyme's backbone and the substrate. Cyan color is lowest RMSf and red color is highest. Values of RMSf range from 0 to 4.46 Å.

The first step transforms the water molecule in a hydroxide group. This transformation allows the hydroxide to have enough strength to attack and bond with the scissile

carbon of the substrate. The nucleophilic attack, in similar mechanisms, is the rate limiting step. In our study the trend seen in other enzymes is also seen here. The second transition state has an energy difference of 18.6 kcal/mol comparing with the reactants and an energy difference of 19.4 kcal/mol in comparison with the previous intermediate. Initially the O_{nuc} is at 1.948 Å from the zinc and at 2.871 Å from the C_{sci} . The O_{pep} of the substrate is being stabilized by a hydrogen bond of the His566 residue and is distant from the zinc at this point (3.799 Å). At the same time that the O_{nuc} is approaching the C_{sci} and distancing itself from the zinc, the O_{pep} is doing the opposite, approaching the zinc atom. At the second transition state (TS2), the O_{nuc} is at 2.037 Å and is now closer to the C_{sci} at a distance of 1.637 Å.

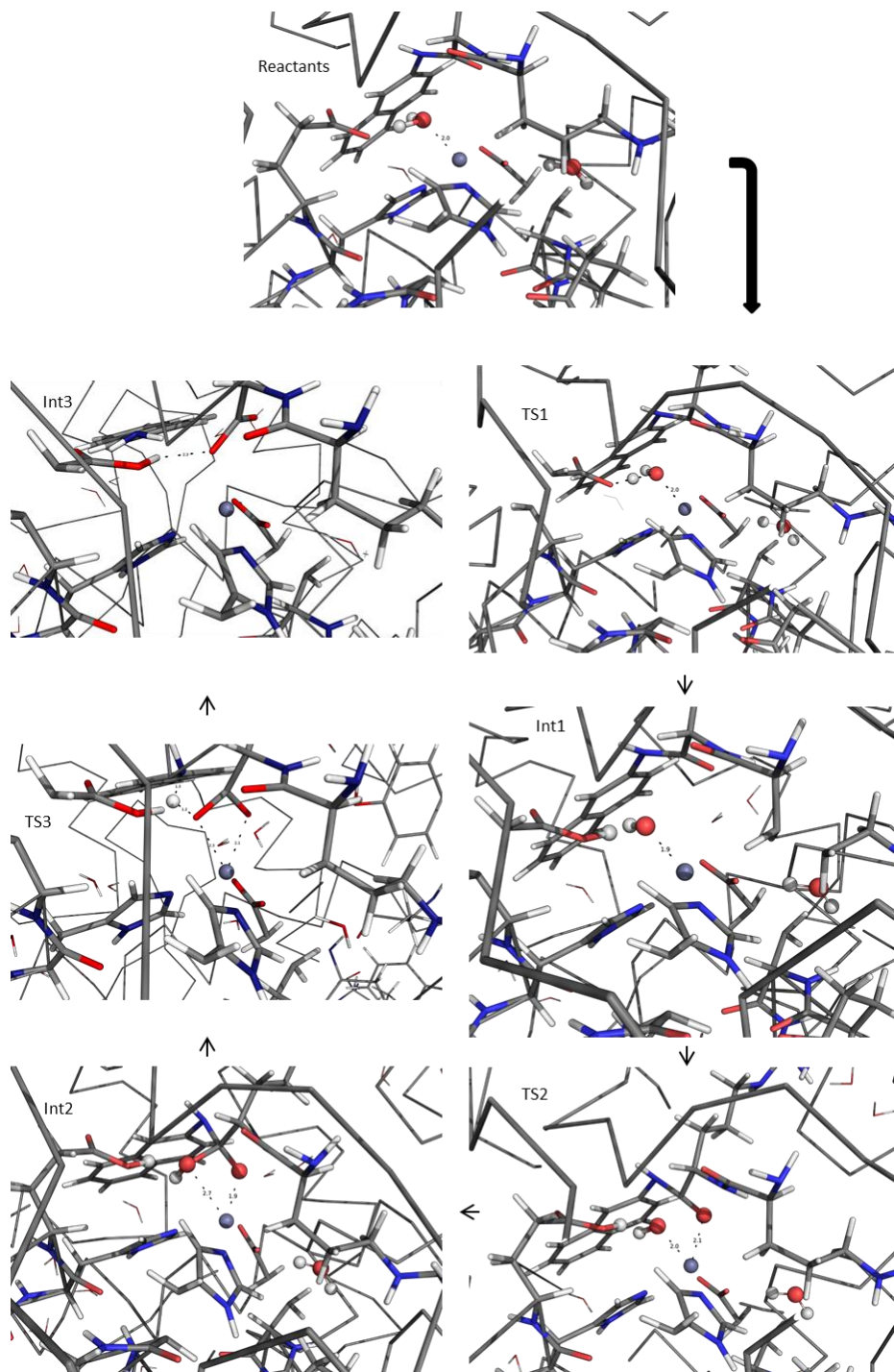


Figure 4- Reaction mechanism obtained for the DPP III.

This approximation the substrate to the O_{nuc} means that the O_{pep} is also closer to the zinc atom. In fact, at the TS2 the distance of the O_{pep} to the zinc atom is now only 2.148 Å, where before it was 3.799 Å. Continuing from the transition state to the second intermediate, the trend continues as before, with the O_{nuc} detaching from the zinc and the O_{pep} binding to zinc, maintaining the former conformation of the zinc.

From the available data from other studies we knew how the reaction should evolve from here. We did a scan from the second intermediate where the O_{nuc} transfers its proton to the amino group on the naphthalene side. The Glu508 residue should also transfer its proton to the O_{nuc} , reprotonating the O_{nuc} . Even though in previous works we saw that this double transfer of protons may occur at the same time, in this study we did not see that happen.

The proton being transfer from the O_{nuc} to the NH-Na is at 1.309 Å from the nitrogen and at 1.186 Å from the O_{nuc} at the TS3 conformation. This transfer has an energetic cost of 6.7 kcal/mol, and the transition state has been confirmed through frequency calculations.

After the transition state is passed and the bonds are formed, the $\text{NH}_2\text{-Na}$ group, distances from the rest of the substrate. The O_{nuc} , being deprotonated, binds again with the zinc atom, while the other oxygen of the carboxyl group unbinds, returning to the conformation seen in the first intermediate.

By protonating the amino group of the naphthalene the bond between nitrogen and carbon no longer exists and the $\text{NH}_2\text{-Na}$ distances itself from the rest of the substrate. The amino group establishes a hydrogen bond $\text{N} - 1.262 \text{ \AA} - \text{H} - 1.730 \text{ \AA} - \text{O}$ (from Glu508) with the Glu508. Because the naphthalene is so big and rigid once it is stabilized, the difference in energy is somewhat higher than we expected, with the minimum after TS3 being 28.8 kcal/mol more stable than the transition state that precedes it.

From this stationary point, the next logic step would be the proton transfer from the Glu508 to the O_{nuc} , however we have not, at this time, completed the calculations for this step.

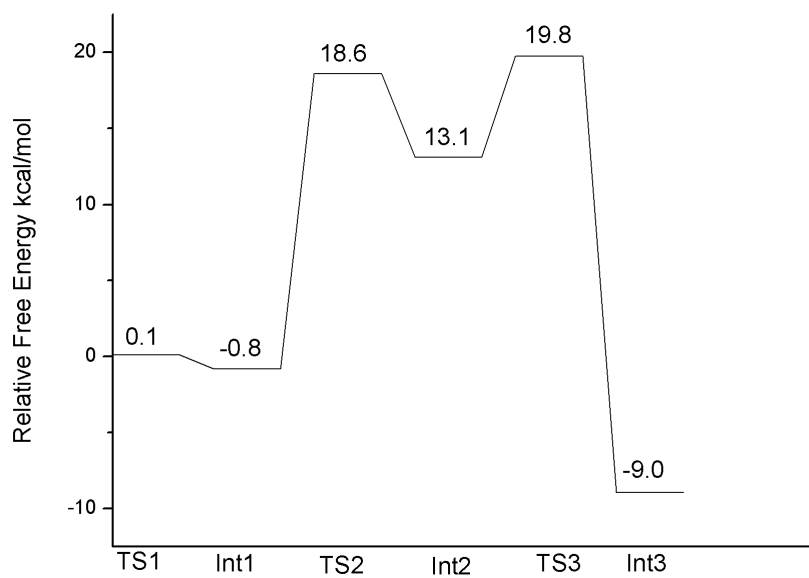


Figure 5- Potential energy surface obtained for the DPP III reaction mechanism obtained using the B3LYP density functional and the 6-311++G(2d,2p) basis set.

4.5 Conclusion

Because of the lack of studies on this enzyme, we believe that this work is important to further our knowledge about the mechanism of the DPP III.

We found in this study a mechanism that is in agreement with the experimental data. Our rate limiting step has as a relative energy of 19.7 kcal/mol while the experimental data tells us that the reaction has an energetic cost around 16.1 kcal/mol.

The DPP III study is not yet completed. We still want to confirm the last step of the reaction mechanism. There may be other questions worth to answer about this mechanism.

As future work we intend to finish the current mechanism and address other possibilities. One of the questions we would like to answer is if the first and second steps that we present here occur at the same time, since the proton transfer of the first step is so energetically cheap.

Another question that interests us is about the intermediate 2. We would like to see if it is possible that instead of the O_{pep} bonding to the zinc when the O_{nuc} breaks its bond, the Glu449 residue would make bond with the free oxygen atom of the carboxyl group.

4.6 Bibliography

1. Musilek, K., Dolezal, M., Gunn-Moore, F. & Kuca, K. Design, evaluation and structure-activity relationship studies of the AChE reactivators against organophosphorus pesticides. *Med. Res. Rev.* **31**, 548–75 (2011).
2. Sturrock, E. D., Natesh, R., van Rooyen, J. M. & Acharya, K. R. Structure of angiotensin I-converting enzyme. *Cell. Mol. Life Sci.* **61**, 2677–86 (2004).
3. Brás, N. F., Fernandes, P. A. & Ramos, M. J. A QM/MM study and MD simulations on the hypertension regulator Angiotensin-Converting Enzyme. (2014).
4. Dive, V., Chang, C.-F., Yiotakis, A. & Sturrock, E. D. Inhibition of zinc metallopeptidases in cardiovascular disease--from unity to trinity, or duality? *Curr. Pharm. Des.* **15**, 3606–21 (2009).
5. Zhang, T. *et al.* Theoretical insights into the functioning of metallopeptidases and their synthetic analogues. *Acc. Chem. Res.* **48**, 192–200 (2015).
6. Pelmeshnikov, V., Blomberg, M. R. A. & Siegbahn, E. M. A theoretical study of the mechanism for peptide hydrolysis by thermolysin. doi:10.1007/s007750100295
7. Gomis-Rüth, F. X., Botelho, T. O. & Bode, W. A standard orientation for metallopeptidases ☆. *BBA - Proteins Proteomics* **1824**, 157–163 (2012).
8. Khaket, T. P., Redhu, D., Dhanda, S. & Singh, J. In Silico Evaluation of Potential DPP-III Inhibitor Precursors from Dietary Proteins. *Int. J. Food Prop.* **18**, 499–507 (2015).
9. Jajč Anin-Jozi, N., Tomi, S. & Abrami, M. Importance of the three basic residues in the vicinity of the zinc-binding motifs for the activity of the yeast dipeptidyl peptidase III. (2013). doi:10.1093/jb/mvt093
10. Fukasawa, K. M., Hirose, J., Hata, T. & Ono, Y. In rat dipeptidyl peptidase III, His568 is essential for catalysis, and Glu507 or Glu512 stabilizes the coordination bond between His455 or His450 and zinc ion. *BBA - Proteins Proteomics* **1804**, 2063–2069 (2010).
11. Hirose, J. *et al.* Characterization of the metal-substituted dipeptidyl peptidase III (rat liver). *Biochemistry* **40**, 11860–5 (2001).
12. Arolas, J. L., Botelho, T. O., Vilcinskis, A. & Xavier Gomis-Rüth, F. Peptide-Bond Synthesis Structural Evidence for Standard-Mechanism Inhibition in Metallopeptidases from a Complex Poised to Resynthesize a Peptide Bond**. doi:10.1002/anie.201103262
13. Bezerra, G. A. *et al.* Entropy-driven binding of opioid peptides induces a large domain motion in human dipeptidyl peptidase III. doi:10.1073/pnas.1118005109
14. Jajčanin - Jozić, N., Deller, S., Pavkov, T., Macheroux, P. & Abramić, M. Identification of the reactive cysteine residues in yeast dipeptidyl peptidase III. *Biochimie* **92**, 89–96 (2010).
15. Chemica, C., Ccaca, A., Jajčanin-Jozić, N., Macheroux, P. & Abramić, M. Yeast Ortholog of Peptidase Family M49: the Role of Invariant Glu. *Croat. Chem. Acta* **85**, 535–540 (2012).

16. Tomi, A. *et al.* Human dipeptidyl peptidase III: insights into ligand binding from a combined experimental and computational approach. doi:10.1002/jmr.1115
17. Tomi, A., Berynskyy, M., Wade, R. C. & Tomi, S. Molecular simulations reveal that the long range fluctuations of human DPP III change upon ligand binding. *Mol. BioSyst. Mol. BioSyst* **11**, 3068–3080 (2015).
18. Bernstein, F. C. *et al.* The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–42 (1977).
19. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–25 (2006).
20. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. D. *et al.* No Title. (2012). at <<http://ambermd.org/>>
21. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. . Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
22. Frisch, M. J. *et al.* Gaussian 09. 2009 (2009). doi:10.1159/000348293
23. Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K. & Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct. THEOCHEM* **461-462**, 1–21 (1999).
24. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
25. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648 (1993).
26. Stephens, P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **98**, 11623–11627 (1994).
27. Vosko, S. H., Wilk, L. & Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **58**, 1200–1211 (1980).
28. Becke, A. D. Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **104**, 1040 (1996).
29. Schmider, H. L. & Becke, A. D. Optimized density functionals from the extended G2 test set. *J. Chem. Phys.* **108**, 9624 (1998).
30. Silvestrelli, P. L. & Parrinello, M. Structural, electronic, and bonding properties of liquid water from first principles. *J. Chem. Phys.* **111**, 3572 (1999).

5 Assessing the promiscuity of the Organo-phosphate degrading enzyme from *Agrobacterium radiobacter*

5.1 Abstract

Organo-phosphate degrading enzyme from *Agrobacterium radiobacter* exhibits promiscuity, not only in the reaction it catalyzes, but also in the metals it uses to catalyze those reactions. Here, we studied three different pairs of binuclear metal centers, Cd – Cd, Zn – Fe and Mn – Mn, for both the hydrolysis of trimethylphosphate and the successive phosphodiester hydrolysis. Both mechanisms have been studied at DFT level of theory using a cluster model approach. For the case of the pair Mn – Mn, calculations were made also with antiferromagnetic coupling, for both high and low spin. For the three different pairs, various spin states were studied to assess the spin state with the lowest initial energy.

The study confirms that the hydrolysis reaction of the phosphotriester is faster than the phosphodiester and in some cases the phosphodiester reaction seems to be too slow to be present in nature.

Our activation energies are in agreement with previous experimental results for the phosphotriesterase mechanism.

5.2 Introduction

Binuclear metallohydrolases such as purple acid phosphatase, urease and the organophosphate degrading enzyme from *agrobacterium radiobacter* have similar active sites with a pair of metal atoms. It may be a pair of homovalent metal atoms, such as the two nickel atoms (Ni^{2+}) of urease, or a pair of heterovalent metal atoms as in the case of purple acid phosphatase, with one conserved iron (Fe^{3+}) atom and a second Zn^{2+} , Fe^{2+} , or Mn^{2+} atom. When a binuclear metallohydrolase has homovalent metal atoms, these may be the same metal or different metal atoms.

In the purple acid phosphatase active site, both metal atoms are pentacoordinated in a trigonal bipyramid way. The iron atom has a first coordination sphere composed of two aspartate residues, one histidine, one tyrosine and one hydroxide. The zinc coordination sphere consists of one asparagine, two histidine residues, and two residues shared with the iron atom: an aspartate and a hydroxide molecule. In the case of urease, which has a very different substrate, we see that the active site is nonetheless very similar. Each nickel atom binds to two histidine residues, one water molecule and bridging the two metal atoms is a hydroxide and a carboxyl group, in this case is a naturally modified lysine residue.

With the revolution of agriculture 70 years ago, the usage of organophosphate esters pesticides grew suddenly. Not only are these organophosphates still used in agriculture but there are also some organophosphate compounds that are used as weapon, such as sarin and VX gases. Nerve agents like sarin are toxic for their ability to hinder acetyl cholinesterase (AChE) normal activity. This enzyme plays a critical role in the normal function of nerve cells. For this reason there is a wide range of organisms that may be endangered in the presence of organophosphate triesters[1], [2].

Besides their use as a weapon, organophosphate compounds are still used to protect crops from insects, leading ultimately to the death of the insects through paralysis. The wide use of these compounds in agriculture means that insects are not the only beings being affected[3], efforts to diminish the usage of these compounds in agriculture are already being made.

The general chemical structure of organophosphate compounds is a phosphate center with three esters bonded and a hydroxide group. Two of the ester groups in the phosphate are fairly stable, while the remaining ester group is more labile, as we shown in this work, and have been proposed before.[4]

There are, however, some strains of bacteria that have shown the ability to hydrolyze organophosphates, which hydrolyze these compounds. *Pseudomonas diminuta*, *Flavobacterium*, *Enterobacter aerogenes* and *Agrobacterium radiobacter* are examples of bacteria with organophosphate degrading enzymes. The enzymes in these bacteria have a sequence identity with over 90% similarity. This similarity shows that different enzymes evolved to reach a similar or the same result[5], [6].

The overall structure of the organophosphate degrading enzyme from *agrobacterium radiobacter* (OPDA) is a homodimer; subunits fold to form a TIM barrel as is typical of amidohydrolases. In OPDA, the binuclear metal center is located in the C-terminus of the β barrel.²⁸ Similar to organophosphate degrading enzyme from *Enterobacter aerogenes* (GpdQ), a combination of crystallographic and magnetic circular dichroism (MCD) structural data indicates the presence of a five- and a six-coordinated metal ions in the active site; however, while in GpdQ the α -site has six ligands and the β -site has five ligands, the situation is reversed in OPDA

The more buried right hand side - site in OPDA (α site) is formed by two histidine residues and an aspartate residue, while the more solvent-exposed β -site also has two histidine ligands in addition to two terminal water molecules. The two metal ions are bridged by a carboxylated lysine and a hydroxide. Disruption of this hydrogen bond network through mutations leads to OPDA forms with catalytic properties and substrate specificities similar to that of phosphotriester hydrolysis (OPH).

Binuclear metallohydrolases can be promiscuous in relation to their metal ions, depending on their environment availability. In the case of *agrobacterium radiobacter*, the most common configuration is Zn-Fe, although previous reports indicate that other configurations can be more efficient. There are some cases in binuclear metallohydrolases where there is some promiscuity when it comes to the metal atoms in the active site.[5], [7], [8] In some cases it depends on the environment availability of the different metals. In the case of *agrobacterium radiobacter*, the subject of our work, it is noted for mostly having an active site with the Zn – Fe combination, nevertheless it was thought previously that other combinations of metals in the active site would be more efficient.[9]–[12]

Besides the promiscuity in the metal atoms of the active site, there may also be promiscuity in the type of catalytic substrate.[13]–[15] The reason for this promiscuity is said to be the fast evolution (evolutionary scale speaking) of organophosphate enzymes caused by the exponential growth in the use of organophosphates in agriculture, as mentioned before.[16] With the growth of the bioavailability of organophosphate compounds, bacteria started using it as a source of phosphorus.

A model of this enzyme with two cobalt atoms in the active site has been previously studied and its mechanism proposed.[17] We thought that further studies of this enzyme with different metal pairings in the active site were needed to fully appreciate the promiscuity of the OPDA catalytic mechanism.

Data from the literature lets us understand how the mechanism should evolve. Scheme 1 shows the computationally studied mechanism for the OPDA with a binuclear cobalt active site. When the substrate enters and binds the active center, the unsubstituted oxygen is orientated to the β -metal and the metal bridging hydroxide moves towards

the α -metal. This new conformation allows for an easier nucleophilic attack of the hydroxide group to the organophosphate substrate.

In the previous work, it is clear that the second hydrolysis mechanism is slower. In our approach, one of the goals was to study the viability of this second mechanism with others metal atoms pairings in the active site. These findings will be discussed further down in this paper. From this previously published paper we also know how the second hydrolysis mechanism works.

At the end of the first mechanism, a molecule of methanol leaves the active site. The conformations of the binuclear metal atoms are mostly the same with the substrate, now a phosphodiester, bonded to both atoms.

Since we are studying metallic active sites we chose to run all geometry optimizations and energy calculations in a model of the active site, instead of using the whole enzyme. Since the active site has two metal atoms, we use the cluster model to ensure that all the activity and chemistry of the enzyme occurs next to this binuclear site. We also chose the residues with great care when modeling the active site, which is a big aspect of obtaining good reproducible results.

Our three models of the active site, the most common binuclear pairing for this enzyme Zn – Fe and other two pairings said to catalyze the same reaction: Mn – Mn and Cd – Cd metal pairings will also be compared with the previously published results.

5.3 Methods

The models were built from the 2D2G[18] structure from the protein databank. The substrate was modelled from a dimethyl thiophosphate to a trimethyl phosphate.

The models obtained were optimized in the gas phase with the B3LYP[19], [20] density functional which includes the Becke's hybrid exchange and correlation functional of Lee, Yang and Parr as implemented in the Gaussian 09 package.[21] The initial model is shown in figure 1.

To find the transition states, we used flexible scans along the reaction pathway. After the geometry of the transition states were roughly known, an optimization was carried out for the transition state. To ensure that the minima found corresponded to same reaction coordinate as the transition state, IRC calculations were made.

The 6-31G(d,p)[22] basis set was used to describe heavy elements and hydrogens and the LANL2DZ[23] pseudopotential was used to describe transition metal ions. Zero point energies were calculated at the same level as geometry optimizations, within the harmonic approximation. Thermal corrections (at 310 K) and entropy were calculated within the ideal gas/rigid rotor/harmonic oscillator model. Further dispersion calculations were carried for the B3LYP functional

Despite the performance of the B3LYP functional in describing the mechanisms involved in enzymatic catalysis supported by the literature, it is known that other functionals, developed during the last decade,[22], [24]–[30] represent a better tool for the calculation of barrier heights, thermochemical kinetics and non-bonded interactions. For this reason, in order to obtain more accurate energies in protein environment, single points calculations were performed on the optimized geometries at the B3LYPD, MPWB1K and M06L levels using the larger basis set 6-311+G(2d,2p) in the framework of the conductor-like self-consistent reaction field polarized continuum model (CPCM).[31]–[36]

A dielectric constant of $\epsilon = 4$, accounting for the average effect of both the protein and the water medium surrounding the protein, was chosen to describe the protein environment. Previous studies, also including charged [17] systems, demonstrated that geometry relaxation effects obtained in solvent are usually rather small. As a consequence, only single-point calculations on optimized structures have been carried out to correct the corresponding free energy for the gas phase by the solvation energy.

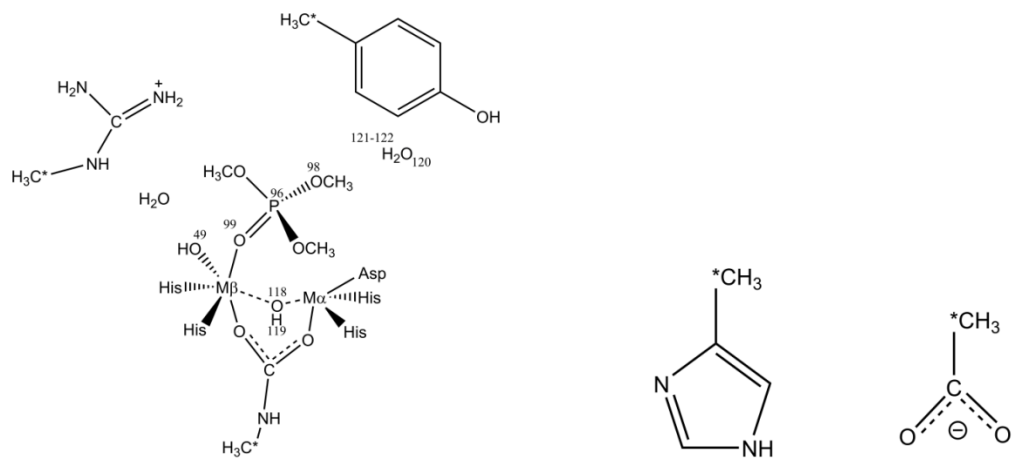


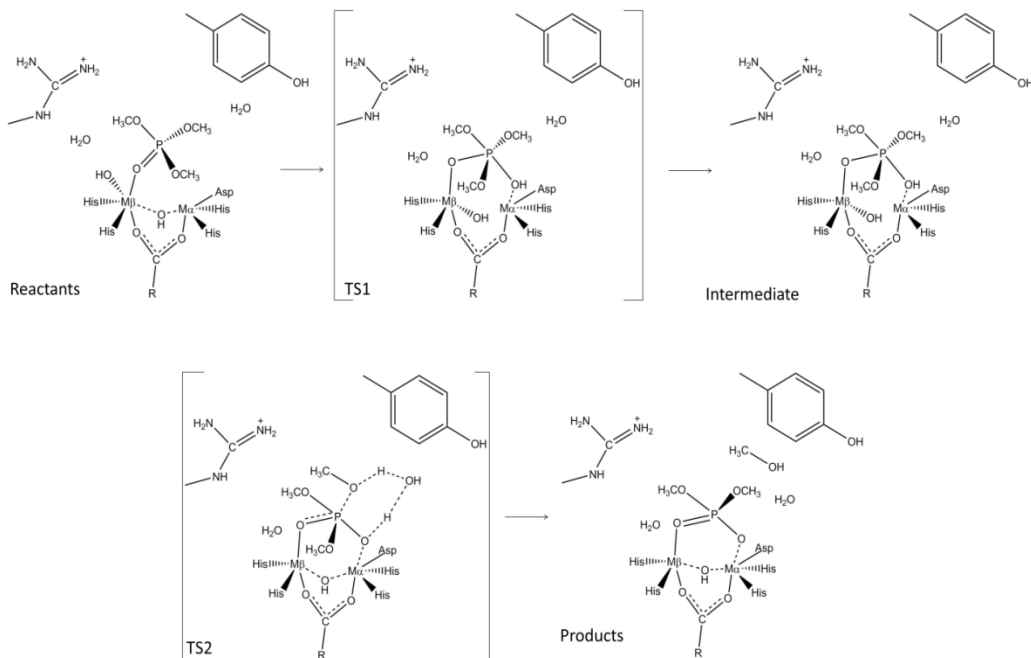
Figure 1- Reactants model with numbered atoms. Stared atoms s where frozen geometry optimizations. Histidine side chain and aspartate side chain on the right hand side to show were the metal ligands were truncated.

5.4 Results

5.4.1 Phosphotriesterase mechanism:

The OPDA mechanism has been studied previously. Nevertheless, since this is a highly promiscuous enzyme and we knew that a change in the metals of the active site would let the enzyme maintain its activity, we decided to study the mechanism computationally for its more common binuclear metal pairing (ZnII and FeII), and other two pairings. The results we obtained are shown and discussed herein.

The reaction mechanisms studied in this work, with three different metal pairs in the active site, are shown in scheme 1 (Trimethyl-phosphatase) and scheme 2 (Dimethyl-phosphatase). The potential energy surfaces and corresponding representations of the optimized structures of minima and transition states are shown subsequently. Important distances and angles are shown in the respective figures. Energies obtained with the different functionals adopted in this study are shown together in one potential energy surface.



Scheme 1 – Proposed mechanism for the Phosphotriesterase reaction of the Organo-phosphate degrading enzyme from *Agrobacterium radiobacter*. M α stands for metal site alpha and M β stand for metal site beta.

Amino acids present in the cluster model are represented with their respective three letter code. R stands for the lysine residue included in the model.

As it was expected, the trimethyl phosphatase reaction mechanism is the one that presents the lowest activation energy. This organophosphate degrading enzyme is first and foremost known as a triester phosphatase enzyme. The reaction mechanism of this enzyme has only recently been described. The reaction mechanism and energies have been published recently for an active site composed of two cobalt atoms. Also, metal ion replacement and magnetic circular dichroism (MCD) studies were published previously which led to the proposal of two different mechanisms. The main difference in these mechanisms lies in the positioning and bonding of the attacking hydroxide group. With the information obtained from the bibliography we began our study with the optimization of the reactants, which in this case means the active site of the OPDA with a trimethyl phosphate bonded to the active site.

In the reactants, the geometry of the metal atoms, resembles a square pyramid for the alpha metal site, and an octahedral for the beta metal site. It is normal in biological environments for the conformations to deviate from the perfect geometry since there are constraints around the active site, imposed by the enzyme itself.

The alpha site is bonded to two histidine residues, one aspartate residue, a lysine residue and the OH group that acts as the nucleophile. On the other hand, the beta site is bonded to two histidine residues, a lysine residue (bridging both metal atoms), the nucleophile, another OH group, and the substrate bonded through the non-substituted oxygen of the phosphate group. Full comparison of the distances obtained after the optimization of these reactants is shown on table 1. We see immediately that the $M\alpha$ and $M\beta$ bonds in the Cd – Cd model are longer (2.190 Å) than in the other two models, Zn – Fe (1.968 Å) and Mn – Mn (2.016 Å). This is expected, since cadmium has a larger radius than the other three metals in this study.

Since cadmium is bulkier than the other three metals in this work it is normal and expectable that the distances of the bonds formed with this metal are longer.

As shown in scheme 1, the first step of the phosphotriesterase reaction is the nucleophilic attack by the bridging hydroxide to the substrate. The OH group approaches the phosphate, maintaining similar distances to both metal atoms in the Cd active site. On the other hand the other two models of the active site, with a Mn – Mn pairing and the Zn – Fe pairing have a different, even if not by much, nucleophilic attack. In these two cases the hydroxide group attacks the substrate, breaking its bond with the metal on the beta site and staying close to metal in the alpha site.

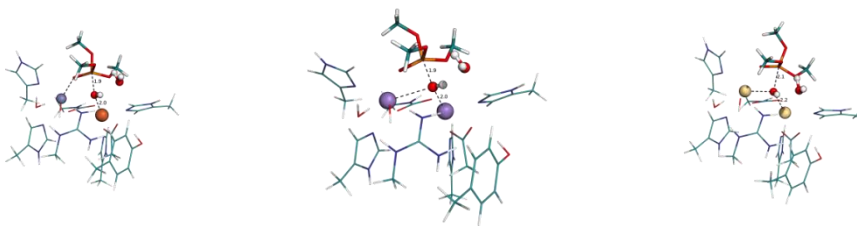


Figure 2- First transition state conformation of the three models. Left to right: Zn – Fe, Mn – Mn and Cd – Cd.

Along the reaction path we see again the differences between the cadmium active site and the other two models. In the reactants of the Cd – Cd model, the nucleophilic attacking group has its proton directed to the water molecule that is going to facilitate the second step. This orientation is maintained on the TS1 conformation. In the other models the hydroxide proton is oriented to an oxygen atom of the substrate. Near transition state 1, the orientation of the proton changes into the orientation seen in the cadmium active site. In transition state 1, we see that the orientation of the proton is consistent to the three models, and even though the distances differ in each case, the proton is clearly forming hydrogen bonds with the close by water molecules in every model.[37]

From the TS1, the OH group starts its energetic way down by continuing the approach to the phosphorus atom of the substrate. The O₁₁₈ closes the distance to the substrate from 1.948 Å at TS1 to 1.678 Å at the intermediate conformation (Zn – Fe), 1.887 Å to 1.670 Å (Mn – Mn) and 2.095 Å to 1.672 Å (Cd – Cd).

At the intermediate, the substrate suffers a rearrangement in its geometry. It is now pentacoordinated with a square pyramid geometry instead of the initial tetrahedral geometry, as seen in figure x. The reorientation of the substrate, and mostly of a particular -OCH₃ group, leads to the repositioning of the water molecule previously stabilizing the nucleophilic group. This water molecule is now stable between the hydroxide group (now bonded to the phosphorus) and the -OCH₃ group. The repositioning of the water molecule is going to enable the second step to occur at lower energy cost than it would be possible without the water molecule working as a proton shuttle.

Moreover, is apparent the new positioning of the hydroxide group, previously bonded only to the metal on the beta site. This group is now acting as a bridge between the alpha site and the beta site, as was the nucleophilic hydroxide at the beginning of the reaction mechanism.

In this intermediate, the geometries of the three models are very close to one another with small differences in the most important distances, as is expected.

Atom1	Atom2	Cd	Mn	ZnFe
<i>Mα</i>	<i>O49</i>	2.190	2.016	1.968
<i>Mα</i>	<i>O118</i>	2.506	2.303	2.217
<i>Mβ</i>	<i>O99</i>	2.192	1.994	2.057
<i>Mβ</i>	<i>O49</i>	2.223	2.001	2.072
<i>Mβ</i>	<i>O118</i>	3.652	3.771	3.534
<i>P96</i>	<i>O98</i>	1.754	1.737	1.759
<i>P96</i>	<i>O118</i>	1.672	1.670	1.678
<i>O98</i>	<i>H121</i>	1.910	1.870	1.900
<i>O118</i>	<i>H119</i>	1.000	1.003	1.006
<i>O120</i>	<i>H119</i>	1.742	1.750	1.703
<i>O120</i>	<i>H121</i>	0.980	0.979	0.980

Table 1- Values from the three models reactant conformation.

The second step of the mechanism is the elongation and breaking of the P–OCH₃ bond facilitated by a water proton shuttle. The leaving group receives a proton from the water molecule, which in turn receives the proton from the OH, now bonded to the substrate. In the transition state 2 conformation, shown in figure x, the leaving group has distanced itself from the substrate and the protons are midway in their path. The bond the oxygen makes with the phosphorus shortens, indicating the leaving of the proton. Since this exchange takes place in a very limited area of the active site model, the general conformation of the model stays as before, as seen in figure 3.

After this transition state, the leaving group gains yet more separation in respect to the substrate and at the products energy minimum the leaving group is at more than 5 Å from the phosphorus atom (figure 3). The water molecule that acted as shuttle remains in place, interacting through an H-bond with the substrate. The substrate itself rearranges its conformation to its initial tetrahedral geometry conformation. Although the conformation is regained, the substrate is now interacting with both metal atoms, where before it was interacting only with the metal at the beta site. We can also see on figure 3 that the initial conformation and bond configuration of the metal atoms are similar to those of the reactants. We see an exchange of the number of bonds each metal is making. In the reactants, the beta site had an octahedral geometry with six bonds and in the products it is establishing five bonds (two histidine residues, one lysine, a hydroxide group and the substrate) and has trigonal bipyramid geometry. On the other hand the metal in the alpha site now has an octahedral geometry configuration (two histidine residues, one aspartate, one lysine, the bridging hydroxide and also the substrate) whereas before it had a square pyramid configuration with five ligands.

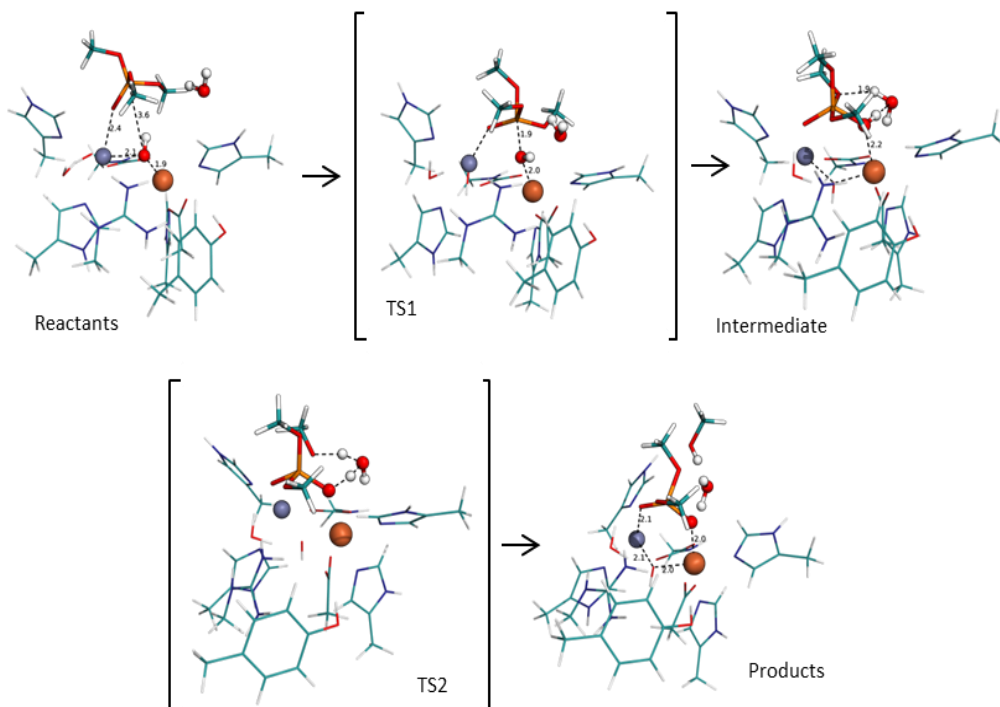


Figure 3 – Obtained mechanism for the phosphotriesterase hydrolysis by the OPDA enzyme with Zn – Fe pairing at the active site.

The potential energy surface of the phosphotriesterase mechanism with MPWB1K is shown on figure 4. The values obtained with the three density functionals for all models are shown in table 2.

As we can see from table 2, the energetics obtained with three different functionals give us different rate determining steps. The B3LYPD functional yields higher activation energy for the first step, i.e., the nucleophilic attack of the hydroxide group, 9.4 kcal/mol compared to 8.4 kcal/mol (Zn – Fe), 9.1 kcal/mol compared to 7.6 kcal/mol (Mn – Mn) and 11.0 kcal/mol compared to 6.6 kcal/mol (Cd – Cd). On the other hand the M06L and MPWB1K density functionals yield higher activation energy for the second mechanistic step, the releasing of the leaving $-OCH_3$ group and concatenated proton shuttle, 8.9 kcal/mol and 7.9 kcal/mol (M06L and MPWB1K) in comparison with 10.9 kcal/mol and 11.2 kcal/mol for the second step (Zn – Fe), 5.1 kcal/mol and 5.2 kcal/mol in comparison with 10.6 kcal/mol and 11.3 kcal/mol (Mn – Mn) and 5.0 kcal/mol and 4.8 kcal/mol in comparison with 7.8 kcal/mol and 9.5 kcal/mol (Cd – Cd).

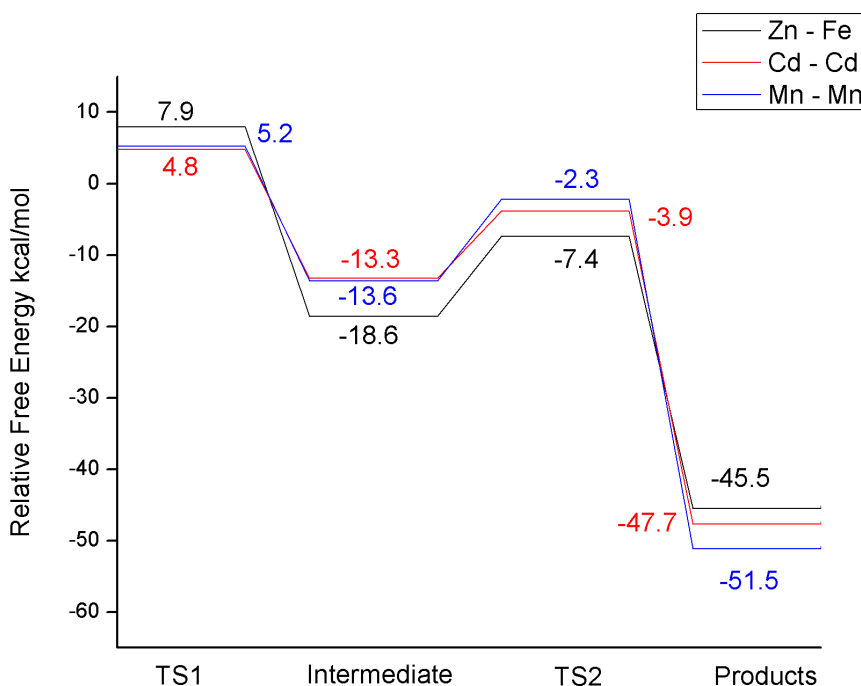


Figure 4- Potential energy surface for the phosphotriesterase hydrolysis by the OPDA enzyme with Zn – Fe pairing at the active site, determined with 6-311++G(2d,2p) basis-sets. Results obtained from single point calculations.

Comparing to the experimental results, obtained for the most common pairing in nature (Zn – Fe), with the energy obtained for our Zn – Fe model, we see that both the M06L and the MPWB1K density functionals yield activation energies close to the obtained experimentally, 12.9 kcal/mol.

The Cd – Cd model differs from the other two models. It has lower activation energies with the M06L and the MPWB1K density functionals for the second step, but the B3LYPD3 density functional yields a higher energy on the first mechanistic step.

The stabilization after the transition states are however similar in all models with every density functional. The Zn – Fe pairing has higher stabilization from the first transition state to the first intermediate, this change on the stabilization after the second transition state where this model is the one that has the lowest stabilization. The Cd – Cd model and the Mn – Mn model, have similar energy differences in the intermediate -13.3 kcal/mol and -13.6 kcal/mol, respectively, in comparison to the reactants.

The energies differences for the products in all three models are very close to one another, -45.5 kcal/mol (Zn – Fe), -51.1 kcal/mol (Mn – Mn) and -47.7 kcal/mol (Cd – Cd) leading us to believe that the overall conformation of the active site models are very stable in these products.

		Relative Free Energy kcal/mol			
		TS1	Intermediate	TS2	Products
Zn – Fe	B3LYPD	9.4	-17.0	-8.6	-51.0
	M06L	8.9	-16.6	-5.7	-48.5
	MPWB1K	7.9	-18.6	-7.4	-45.5
Mn – Mn	B3LYPD	9.1	-7.6	0.0	-47.4
	M06L	5.1	-10.9	-0.3	-48.4
	MPWB1K	5.2	-13.6	-2.3	-51.1
Cd – Cd	B3LYPD	11.0	-6.2	0.4	-43.0
	M06L	5.0	-9.2	-1.4	-45.4
	MPWB1K	4.8	-13.3	-3.9	-47.7

Table 2- Relative free energy values for the three models with the three density functionals used for the single points calculations. Basis set used was the 6-311++G(2d,2p). Values in bold and italic are represented in figure 6.

5.4.2 Promiscuous phosphodiesterase mechanism:

The products conformations of the active site in the previous reaction mechanism are similar to the reactants conformation of this second mechanism. However, because of the leaving group and its stabilizing water molecule are no longer in the model, the substrate binds more thoroughly on both metal atoms, particularly for the metal in the alpha site. The water molecule that was stabilizing the leaving group was also making an H-bond with the oxygen atom of the substrate. With this H-bond the oxygen shortens its distance to the metal on the alpha site, as seen in figure 5.

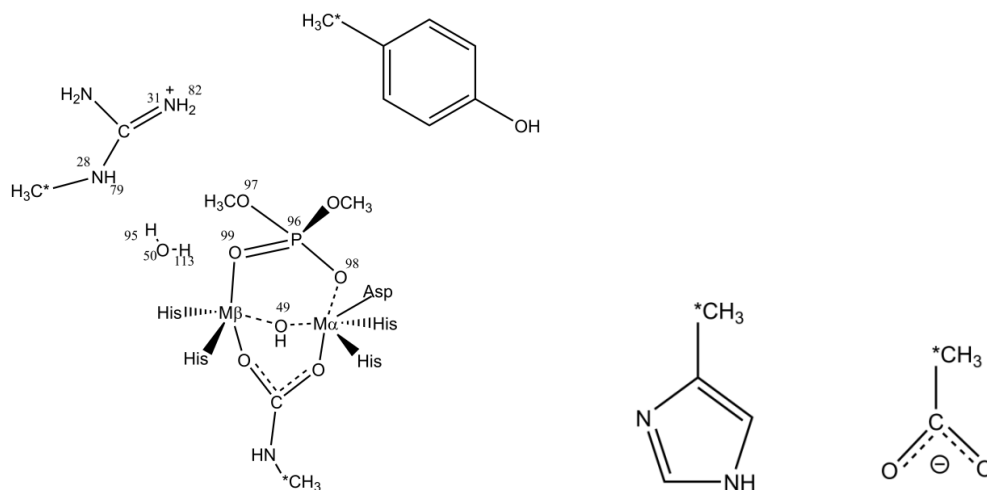
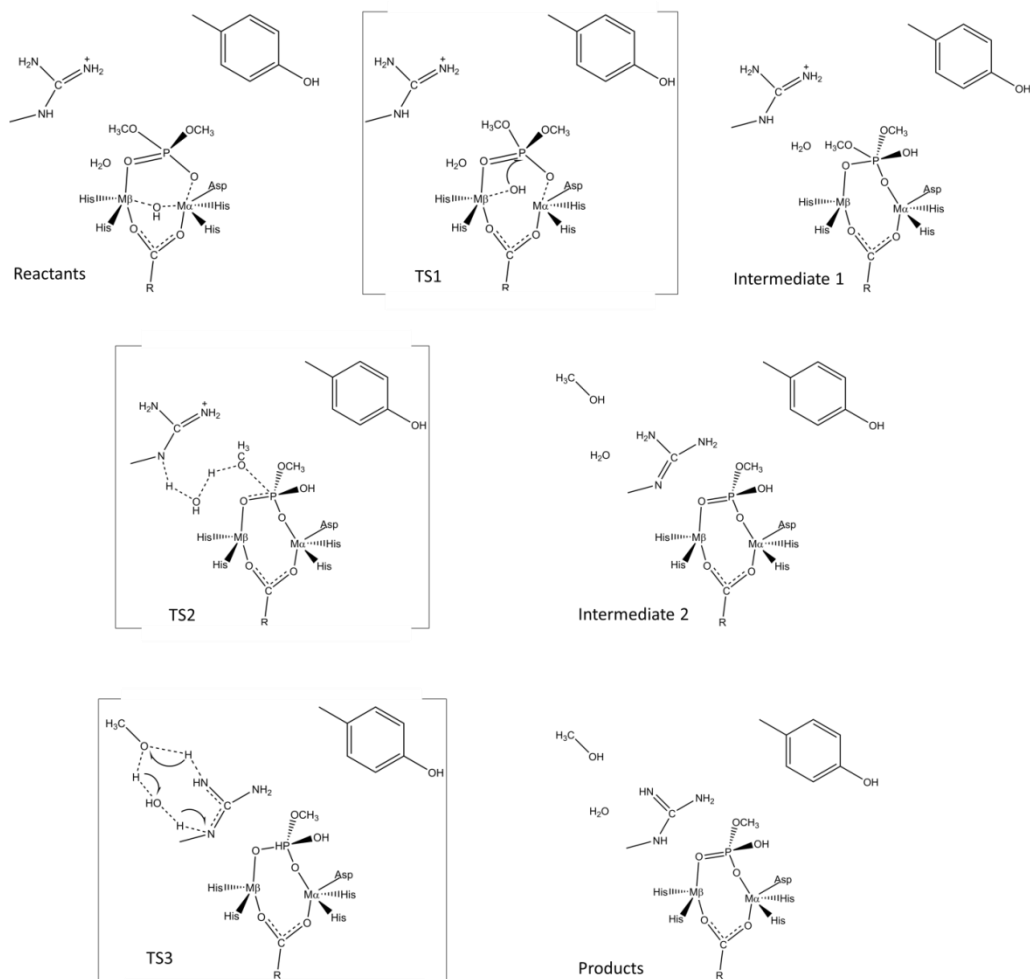


Figure 5- Initial conformation for the hydrolysis of a phosphodiester substrate. Stared numbers were frozen in the geometry optimizations calculations. Residues on the right side, show how the metal ligands were used and frozen.

In the Cd – Cd model, the geometry is not the same as the other two models (supporting information). The metal in the alpha site is not as tightly bonded to the substrate, and the hydroxide is maintained on a bridging position, whereas the hydroxide groups of the other models are clearly bonded only to the metal on the beta site. As a consequence, the hydroxide group is far more stable in this position and the nucleophilic attack on the substrate is more expensive on this model.



Scheme 2 – Proposed mechanism for the Phosphodiesterase reaction of the Organo-phosphate degrading enzyme from *Agrobacterium radiobacter*. M α stands for metal site alpha and M β stand for metal site beta. Amino acids present in the cluster model are represented with their respective three letter code. R stands for the lysine residue included in the model.

The differences on the reactants conformation lead us to predict differences on the attack and the transition state itself. In fact we did find some differences on the transition state conformation, but the attack and the movement the nucleophilic group does towards the transition state has some differences. The main similarity, though, is that, in the three models, the nucleophilic group when starting the approach to the substrate turns the proton to the aspartate residue bonded to the metal on the alpha site for stabilization. However, because the distances to the aspartate residue were not the

same, the hydroxide group of the Zn – Fe model has to travel further than the group on the Mn – Mn model. This forces the nucleophile to approach the alpha site and to act as a bridge between metals on the first transition state whereas on the Mn – Mn model, the nucleophilic group stays bonded only to the beta site metal.

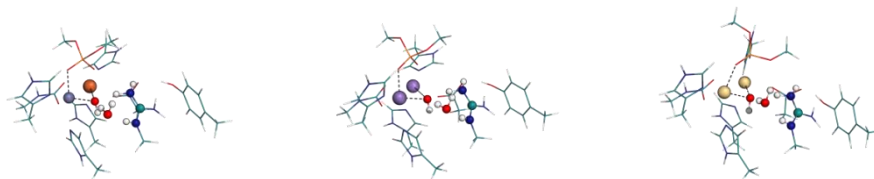


Figure x- Obtained conformation for the three models reactants. From left to right: Zn – Fe, Mn – Mn and Cd – Cd.

On the other hand, as figure x shows, the Cd – Cd model's nucleophilic group distances itself from the alpha site metal when stabilizing with aspartate residue.

Distances from both the reactants and transition states are shown on table x in order to better see the changes that each model undergoes.

Continuing the nucleophilic attack on the substrate, the hydroxide group comes increasingly closer to it, until it bonds and reaches a minimum in energy, the first intermediate of the second reaction mechanism.

The substrate, again, has to adjust to this new ligand as in the first reaction mechanism, going from a tetrahedral geometry to a trigonal bipyramid. The hydroxide group is no longer bonded nor is interacting with either one of the metals, and the substrate, which was bonded to both metal atoms with two oxygen atoms now making a bridge between both metal sites with only one oxygen atom, while the other is only interacting with the metal on the beta site. The water molecule that was stabilizing the hydroxide group in the reactants conformation is now establishing a hydrogen network with the positively charged arginine residue and the substrate (figure 6). This network is important because it helps the exchange of protons among the substrate and the enzyme.[38]

The second step of this mechanism is the release of a leaving $-\text{OCH}_3$ group from the substrate, much as it was in the first mechanism. The water molecule close to leaving group is facilitating this step; it releases a proton to the leaving group and takes another from the arginine. The P – OCH_3 bond is elongated from 1.726 Å in the intermediate to 1.918 Å in the transition state, on the Zn – Fe model, showing the beginning of the separation of the leaving group. In the other two models the elongation is very similar, 1.724 Å to 1.932 Å in the Mn - Mn model and 1.733 Å to 1.918 Å in the Cd – Cd model. The proton from the water is about midway to the $-\text{OCH}_3$ group at 1.273 Å (Zn

– Fe model) and the proton from the arginine is also in the path to re-establish the water molecule 1.155 Å (Zn – Fe model).

Overcoming the second transition state leads to a rearranging of the substrate, the inverse of what happened in the first step, and to the full detachment of the leaving group. The water molecule keeps the H-bond network with the leaving group, the arginine, and the substrate, stabilizing this intermediate a great deal in relation to its previous transition state, 29 kcal/mol in the Zn – Fe model.

Arginine usually becomes neutral by losing a proton of one of the terminal nitrogen atoms. In this case, however, the arginine residue releases the proton of the nitrogen in the middle of the side chain. The reprotonation can as well occur after the releasing of the substrate.

The last step in this promiscuous enzyme is the exchange of the protonation on the arginine residue. On this step there is a three way proton shuttle. The leaving group releases its proton to the arginine previously unprotonated nitrogen. When releasing the proton, the leaving group receives a proton from the water molecule, which in turn receives the proton from the closest –NH₂ of the arginine. The final result of this step is energetically very similar to the second intermediate, only a little more stable.

In these products the leaving group is reprotonated, as is the water molecule. The arginine is neutrally charged with a double bond C=NH in the end of the side chain.

All three models have similar conformations in this second mechanism. We believe that the difference in the metal atom pairing is most important when the chemistry of the reaction occur directly connected to one or both metal atoms.

The energetics of this second mechanism (figure 6) shows us that the promiscuity of the enzyme has its limits. In the case of the Cd – Cd model, the mechanism would be endergonic and also very difficult to overcome with such highly energetic transition states. Having analyzed the phosphotriesterase mechanism, we expected to see higher energetic values with this active site model, however the differences are clear in the case of the phosphodiesterase mechanism. The Mn – Mn model has the highest activation energy of the reaction in first transition state. The transition states energies in this model are progressively lower, from TS1 to TS3 with the last transition state being the easiest to overcome.

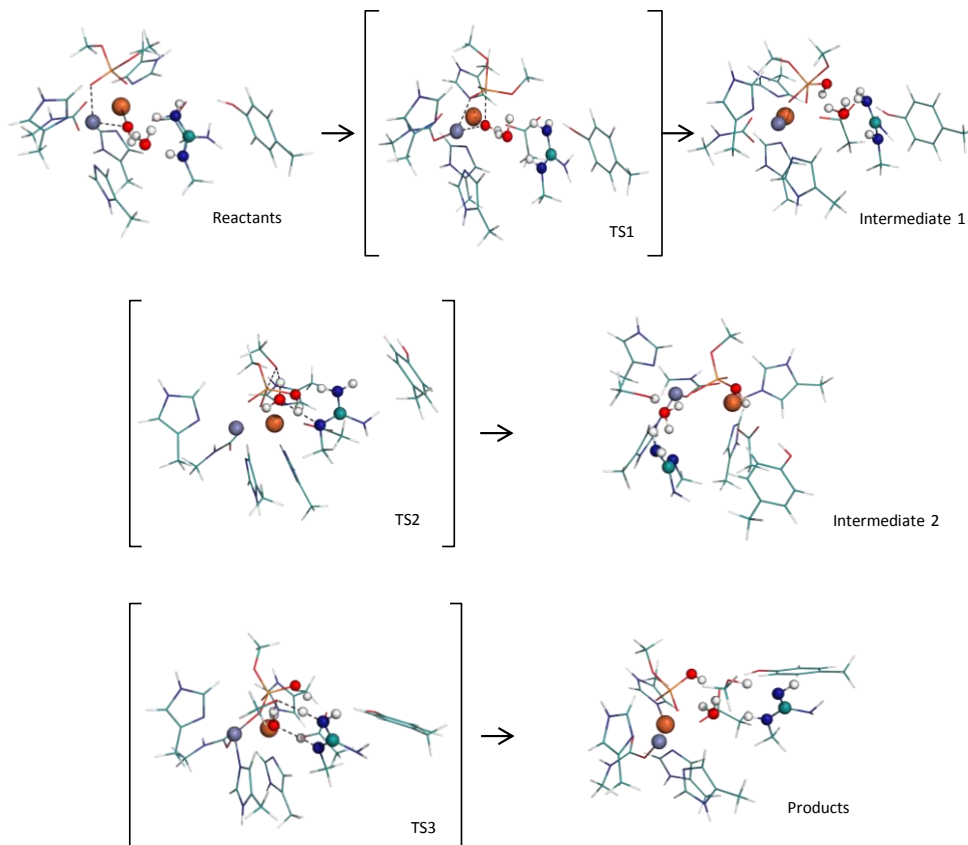


Figure 6 – Obtained mechanism for the phosphodiesterase hydrolysis by the OPDA enzyme with Zn – Fe pairing at the active site.

Finally the Zn – Fe model, being the most common to occur naturally, it was expected to have the best energies of the three models. This however was not confirmed owing to the fact that the second mechanistic step in this model has higher activation energy than any mechanistic step in the Mn – Mn model. Nevertheless it is difficult to indicate a rate limiting step for the phosphodiesterase mechanism since different models yield different results. The highest relative free energy that both the Mn – Mn and Zn – Fe models have to overcome is circa 30 kcal/mol in both cases.

In this mechanism the three density functionals used are in agreement with each other, not only for each mechanistic step, but also for each active site model studied.

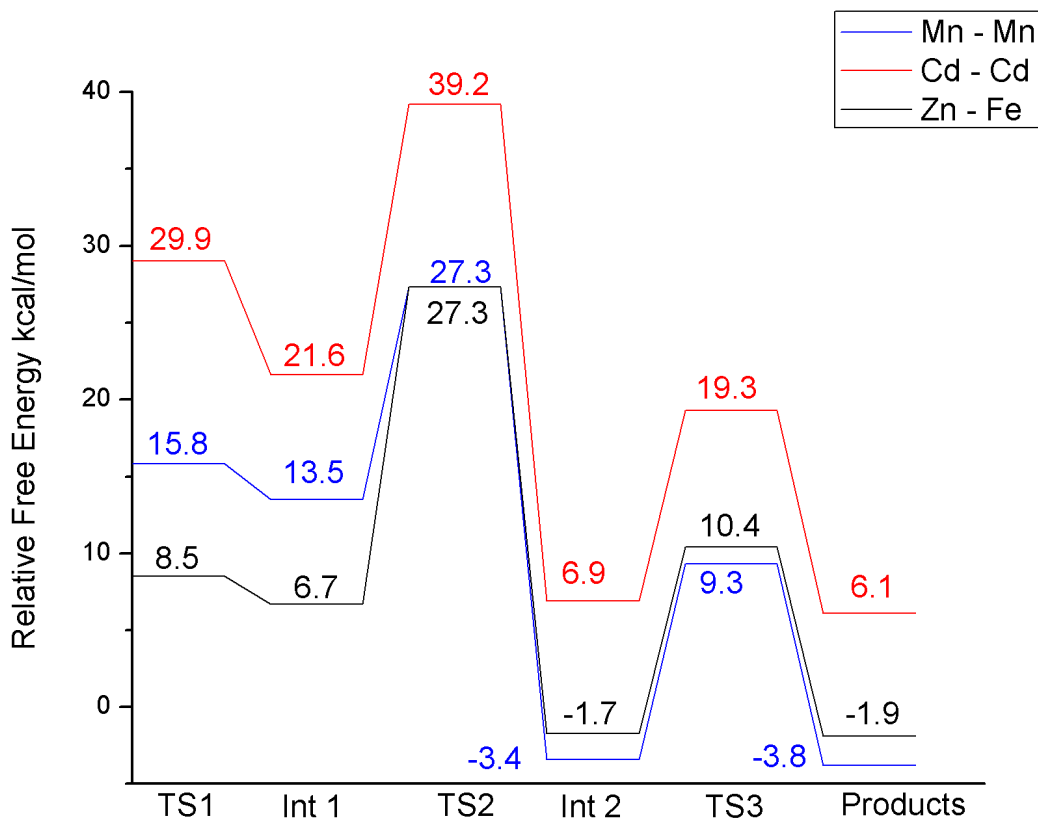


Figure 7- Potential energy surface for the phosphodiesterase hydrolysis by the OPDA enzyme with Zn – Fe pairing at the active site, determined with 6-311++G(2d,2p) basis-sets. Results obtained from single point calculations.

The obtained mechanism for this second reaction is shown in figure 6, as an example of the reaction obtained the model of the active site with the Zn – Fe pairing is shown. A water molecule is once again the facilitator of the proton exchanges that occur along the reaction mechanism, as it happened for the first studied mechanism.

Conclusions

We confirmed in this study the promiscuity of the organophosphate degrading enzyme from *Agrobacterium radiobacter*, not only in the metals this enzyme uses in its active site but also in the type of reaction it catalyzes. We feel that this promiscuity is not complete since there are some metal combinations that catalyze the main reaction (phosphotriesterase mechanism) but do not catalyze the phosphodiesterase mechanism. The results we obtained for the first mechanism are in agreement with the experimental results, obtained with a Zn – Fe pairing in the active site. As it had been proposed other pairings yield better results. This is in agreement with the statement by [Schenk et al.] that the Zn – Fe pairing may be the most common for bioavailability reasons and not for its better ability to catalyze the reaction.

Not having experimental kinetic studies for the second mechanism and the fact that different models have different rate limiting steps we cannot propose a rate limiting step. The highest difference in relative free energy is however in agreement with a previous study of this enzyme and mechanism.[17]

Even though we studied antiferromagnetic coupling in the model with Mn – Mn pairing in the active site (see Supporting Information) it did not contribute to stabilize the initial conformation, this is also in agreement with previous study.

5.5 Bibliography

- [1] K. Musilek, M. Dolezal, F. Gunn-Moore, and K. Kuca, "Design, evaluation and structure-activity relationship studies of the AChE reactivators against organophosphorus pesticides.," *Med. Res. Rev.*, vol. 31, no. 4, pp. 548–75, Jul. 2011.
- [2] F. M. Raushel, "Bacterial detoxification of organophosphate nerve agents," *Curr. Opin. Microbiol.*, vol. 5, no. 3, pp. 288–295, Jun. 2002.
- [3] N. Munro, "Toxicity of the organophosphate chemical warfare agents GA, GB, and VX: implications for public protection.," *Environ. Health Perspect.*, vol. 102, no. 1, pp. 18–38, Jan. 1994.
- [4] A. N. Bigley and F. M. Raushel, "Catalytic mechanisms for phosphotriesterases.," *Biochim. Biophys. Acta*, vol. 1834, no. 1, pp. 443–53, Jan. 2013.
- [5] G. Schenk, N. Mitić, L. R. Gahan, D. L. Ollis, R. P. McGeary, and L. W. Guddat, "Binuclear metallohydrolases: complex mechanistic strategies for a simple chemical reaction.," *Acc. Chem. Res.*, vol. 45, no. 9, pp. 1593–603, Sep. 2012.
- [6] N. Mitić, S. J. Smith, A. Neves, L. W. Guddat, L. R. Gahan, and G. Schenk, "The catalytic mechanisms of binuclear metallohydrolases.," *Chem. Rev.*, vol. 106, no. 8, pp. 3338–63, Aug. 2006.
- [7] N. Mitić, M. Miraula, C. Selleck, K. S. Hadler, E. Uribe, M. M. Pedroso, and G. Schenk, "Catalytic mechanisms of metallohydrolases containing two metal ions.," *Adv. Protein Chem. Struct. Biol.*, vol. 97, pp. 49–81, Jan. 2014.
- [8] D. E. Wilcox, "Binuclear Metallohydrolases.," *Chem. Rev.*, vol. 96, no. 7, pp. 2435–2458, Nov. 1996.
- [9] E. Ghanem, Y. Li, C. Xu, and F. M. Raushel, "Characterization of a phosphodiesterase capable of hydrolyzing EA 2192, the most toxic degradation product of the nerve agent VX.," *Biochemistry*, vol. 46, no. 31, pp. 9032–40, Aug. 2007.
- [10] I. Horne, X. Qiu, D. L. Ollis, R. J. Russell, and J. G. Oakeshott, "Functional effects of amino acid substitutions within the large binding pocket of the phosphotriesterase OpdA from *Agrobacterium* sp. P230.," *FEMS Microbiol. Lett.*, vol. 259, no. 2, pp. 187–94, Jun. 2006.
- [11] K. S. Hadler, E. A. Tanifum, S. H.-C. Yip, N. Mitić, L. W. Guddat, C. J. Jackson, L. R. Gahan, K. Nguyen, P. D. Carr, D. L. Ollis, A. C. Henge, J. A. Larrabee, and G. Schenk, "Substrate-promoted formation of a catalytically competent binuclear center and regulation of reactivity in a glycerophosphodiesterase from *Enterobacter aerogenes*.," *J. Am. Chem. Soc.*, vol. 130, no. 43, pp. 14129–38, Oct. 2008.
- [12] F. Ely, K. S. Hadler, L. R. Gahan, L. W. Guddat, D. L. Ollis, and G. Schenk, "The organophosphate-degrading enzyme from *Agrobacterium radiobacter* displays mechanistic flexibility for catalysis.," *Biochem. J.*, vol. 432, no. 3, pp. 565–73, Dec. 2010.
- [13] C. J. Jackson, P. D. Carr, H.-K. Kim, J.-W. Liu, P. Herrald, N. Mitić, G. Schenk, C. A. Smith, and D. L. Ollis, "Anomalous scattering analysis of

- Agrobacterium radiobacter phosphotriesterase: the prominent role of iron in the heterobinuclear active site.," *Biochem. J.*, vol. 397, no. 3, pp. 501–8, Aug. 2006.
- [14] C. J. Jackson, J.-L. Foo, H.-K. Kim, P. D. Carr, J.-W. Liu, G. Salem, and D. L. Ollis, "In crystallo capture of a Michaelis complex and product-binding modes of a bacterial phosphotriesterase.," *J. Mol. Biol.*, vol. 375, no. 5, pp. 1189–96, Feb. 2008.
- [15] F. Ely, M. M. Pedroso, L. R. Gahan, D. L. Ollis, L. W. Guddat, and G. Schenk, "Phosphate-bound structure of an organophosphate-degrading enzyme from *Agrobacterium radiobacter*.," *J. Inorg. Biochem.*, vol. 106, no. 1, pp. 19–22, Jan. 2012.
- [16] F. Ely, J. L. Foo, C. J. Jackson, L. R. Gahan, D. L. Ollis, and G. Schenk, "Enzymatic bioremediation: Organophosphate degradation by binuclear metallo-hydrolases," *Curr. Top. Biochem. Res.*, vol. 2008, no. 9, pp. 63–78, 2008.
- [17] M. E. Alberto, G. Pinto, N. Russo, and M. Toscano, "Triesterase and promiscuous diesterase activities of a di-Co(II)-containing organophosphate degrading enzyme reaction mechanisms.," *Chemistry*, vol. 21, no. 9, pp. 3736–45, Feb. 2015.
- [18] C. Jackson, H.-K. Kim, P. D. Carr, J.-W. Liu, and D. L. Ollis, "The structure of an enzyme-product complex reveals the critical role of a terminal hydroxide nucleophile in the bacterial phosphotriesterase mechanism.," *Biochim. Biophys. Acta*, vol. 1752, no. 1, pp. 56–64, Aug. 2005.
- [19] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *J. Chem. Phys.*, vol. 98, no. 7, p. 5648, Apr. 1993.
- [20] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, no. 2, pp. 785–789, Jan. 1988.
- [21] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, and D. J. Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, "Gaussian 09." Gaussian, Inc., Wallingford, CT, USA, p. 2009, 2009.
- [22] R. Ditchfield, "Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules," *J. Chem. Phys.*, vol. 54, no. 2, p. 724, Sep. 1971.
- [23] T. H. Dunning Jr and P. J. Hay, "Modern Theoretical Chemistry," in *Modern Theoretical Chemistry*, H. F. Schaefer, Ed. New York: Plenum, 1977, pp. 1–28.
- [24] M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, and P. E. M. Siegbahn, "Quantum chemical studies of mechanisms for metalloenzymes.," *Chem. Rev.*,

- vol. 114, no. 7, pp. 3601–58, Apr. 2014.
- [25] P. E. M. Siegbahn and F. Himo, “Recent developments of the quantum chemical cluster approach for modeling enzyme reactions.,” *J. Biol. Inorg. Chem.*, vol. 14, no. 5, pp. 643–51, Jun. 2009.
- [26] T. Lind, P. E. M. Siegbahn, and R. H. Crabtree, “A Quantum Chemical Study of the Mechanism of Tyrosinase,” *J. Phys. Chem. B*, vol. 103, no. 7, pp. 1193–1202, Feb. 1999.
- [27] P. E. M. Siegbahn and M. R. A. Blomberg, “Transition-Metal Systems in Biochemistry Studied by High-Accuracy Quantum Chemical Methods,” *Chem. Rev.*, vol. 100, no. 2, pp. 421–438, Feb. 2000.
- [28] L. Noodleman, T. Lovell, W.-G. Han, J. Li, and F. Himo, “Quantum chemical studies of intermediates and reaction pathways in selected enzymes and catalytic synthetic systems,” *Chem. Rev.*, vol. 104, no. 2, pp. 459–508, Feb. 2004.
- [29] R. Ditchfield, “Self-consistent perturbation theory of diamagnetism,” *Mol. Phys.*, vol. 27, no. 4, pp. 789–807, Apr. 1974.
- [30] M. S. Gordon, “The isomers of silacyclopropane,” *Chem. Phys. Lett.*, vol. 76, no. 1, pp. 163–168, Nov. 1980.
- [31] M. E. Alberto, T. Marino, M. J. Ramos, and N. Russo, “Atomistic details of the Catalytic Mechanism of Fe(III)–Zn(II) Purple Acid Phosphatase,” *J. Chem. Theory Comput.*, vol. 6, no. 8, pp. 2424–2433, Aug. 2010.
- [32] S. F. Sousa, P. A. Fernandes, and M. J. Ramos, “General performance of density functionals.,” *J. Phys. Chem. A*, vol. 111, no. 42, pp. 10439–52, Oct. 2007.
- [33] A. J. M. Ribeiro, M. J. Ramos, and P. A. Fernandes, “Benchmarking of DFT Functionals for the Hydrolysis of Phosphodiester Bonds,” *J. Chem. Theory Comput.*, vol. 6, no. 8, pp. 2281–2292, Aug. 2010.
- [34] A. J. M. Ribeiro, M. E. Alberto, M. J. Ramos, P. A. Fernandes, and N. Russo, “The catalytic mechanism of protein phosphatase 5 established by DFT calculations.,” *Chemistry*, vol. 19, no. 42, pp. 14081–9, Oct. 2013.
- [35] M. E. Alberto, T. Marino, N. Russo, E. Sicilia, and M. Toscano, “The performance of density functional based methods in the description of selected biological systems and processes.,” *Phys. Chem. Chem. Phys.*, vol. 14, no. 43, pp. 14943–53, Nov. 2012.
- [36] T. Marino, N. Russo, and M. Toscano, “What occurs by replacing Mn²⁺ with Co²⁺ in human arginase I: first-principles computational analysis.,” *Inorg. Chem.*, vol. 52, no. 2, pp. 655–9, Jan. 2013.
- [37] G. A. Jeffrey, *An Introduction to Hydrogen Bonding* | George A. Jeffrey | 9780195095494 | Oxford University Press Canada. Oxford: Oxford University Press, 1997.
- [38] G. P. Pinto, N. F. Brás, M. A. S. Perez, P. A. Fernandes, N. Russo, M. J. Ramos, and M. Toscano, “Establishing the Catalytic Mechanism of Human Pancreatic α -Amylase with QM/MM Methods.,” *J. Chem. Theory Comput.*, vol. 11, no. 6, pp. 2508–16, Jun. 2015.

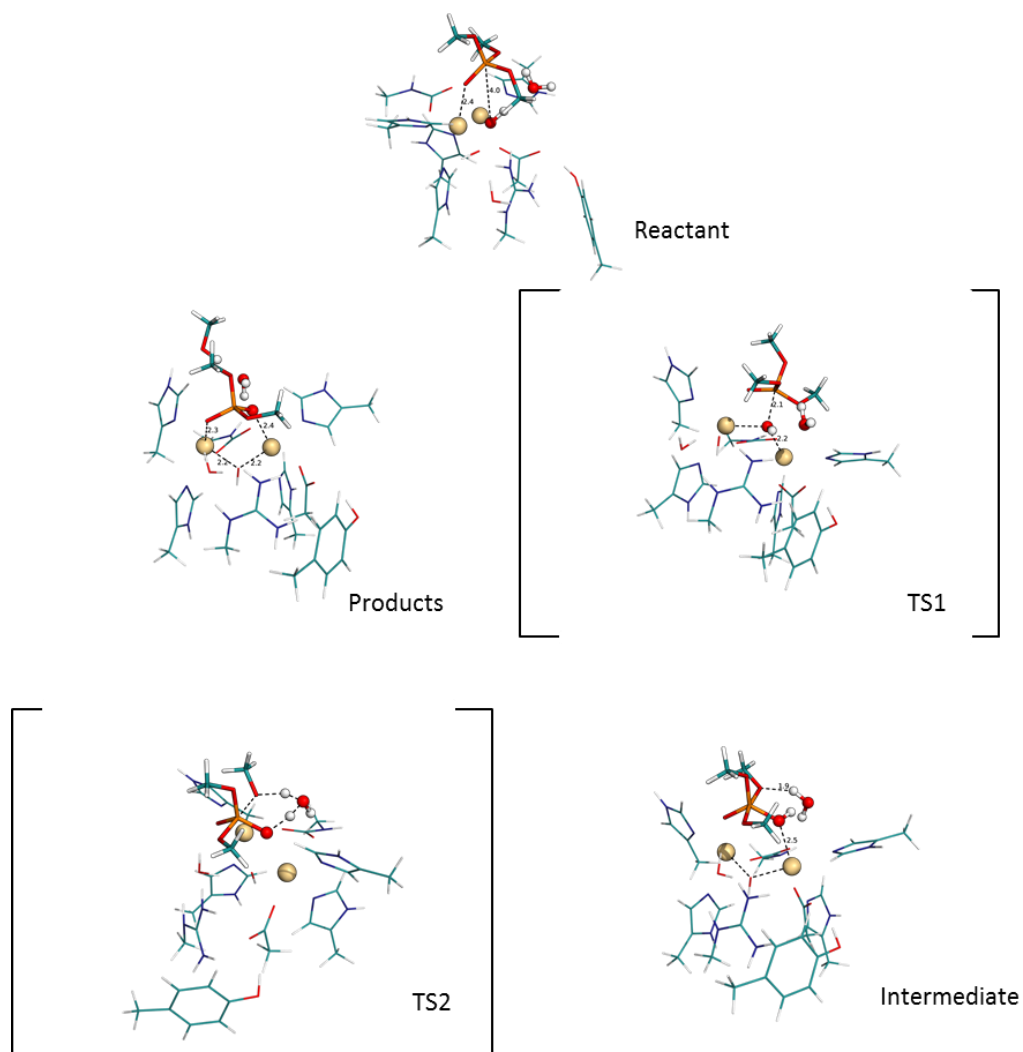


Figure 2- Obtained mechanism for the Cd – Cd model. Phosphotriesterase mechanism

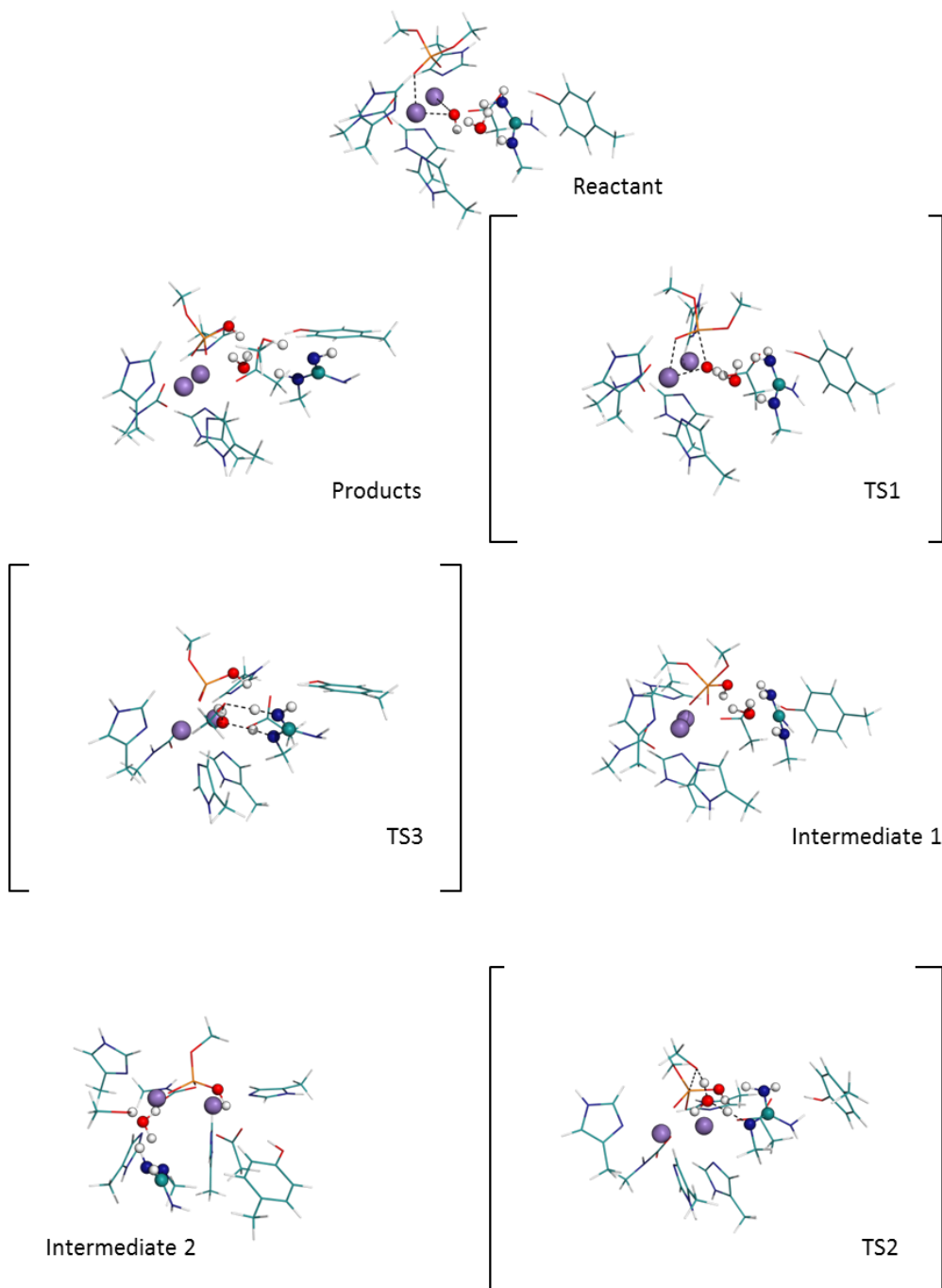


Figure 3- Obtained mechanism for the Mn – Mn model. Phosphodiesterase mechanism

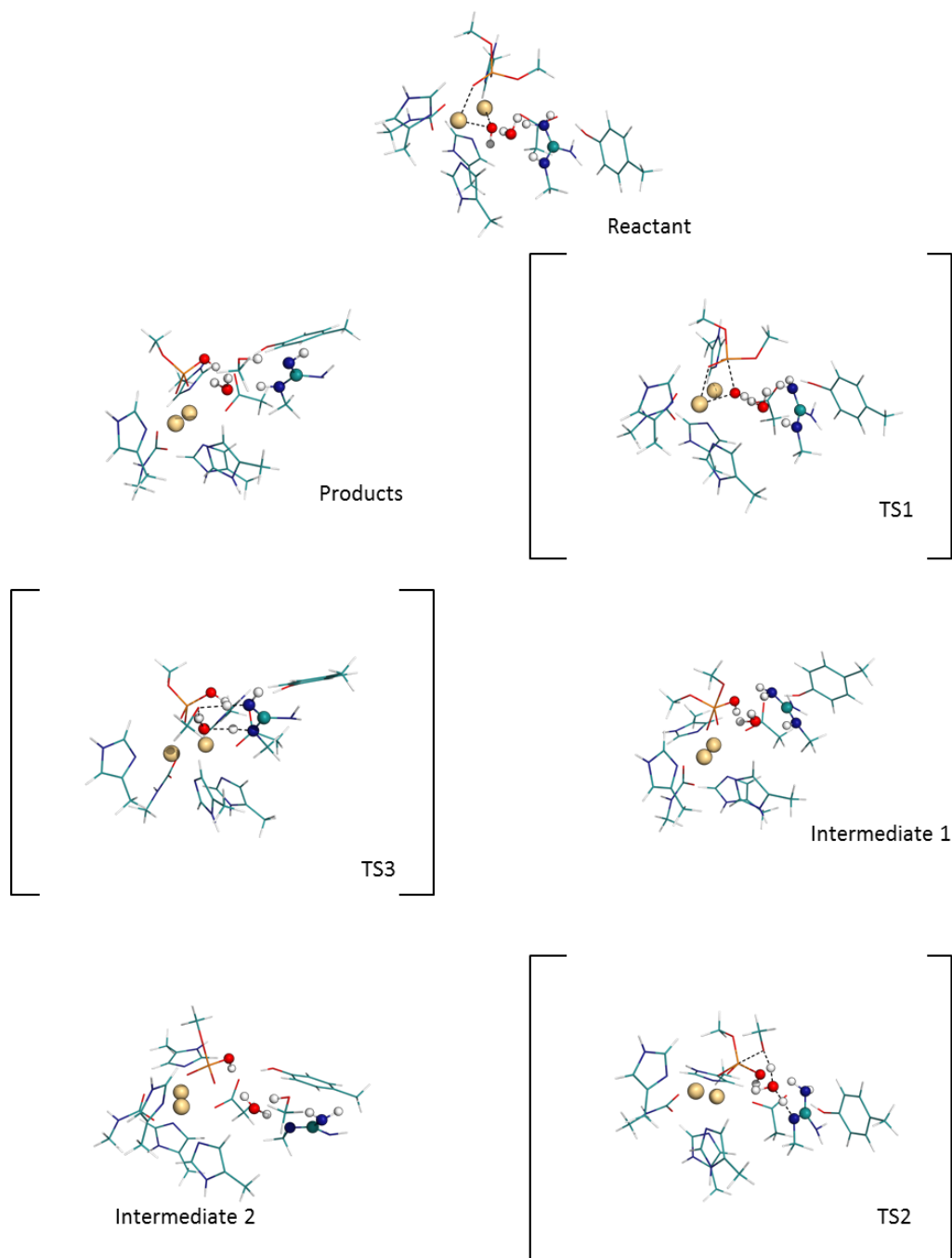


Figure 4- Obtained mechanism for the Cd – Cd model. Phosphodiesterase mechanism

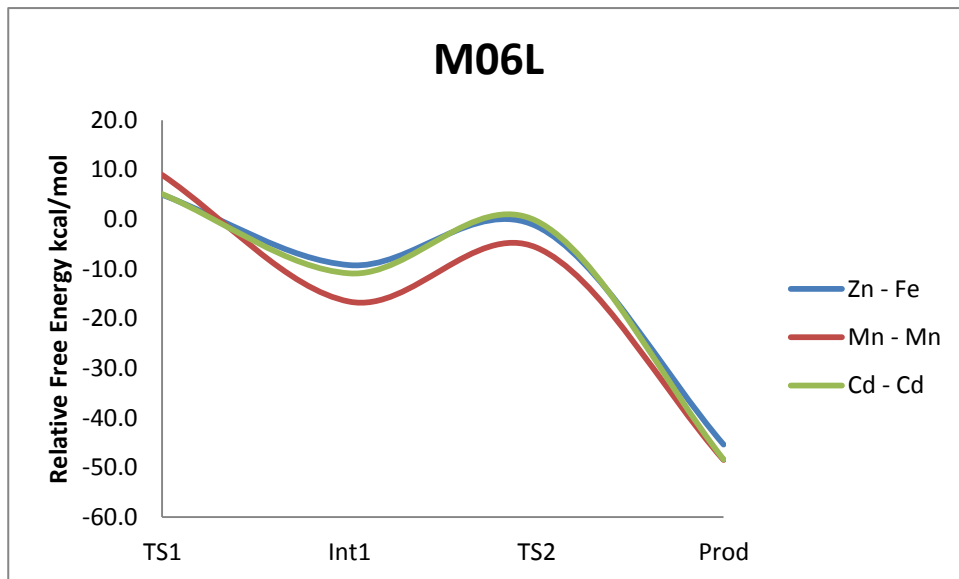


Figure 5- Potential energy surface obtained for the three models with the M06L density functional and 6-311++G (2d,2p) basis-set. Values for the phosphotriesterase reaction mechanism.

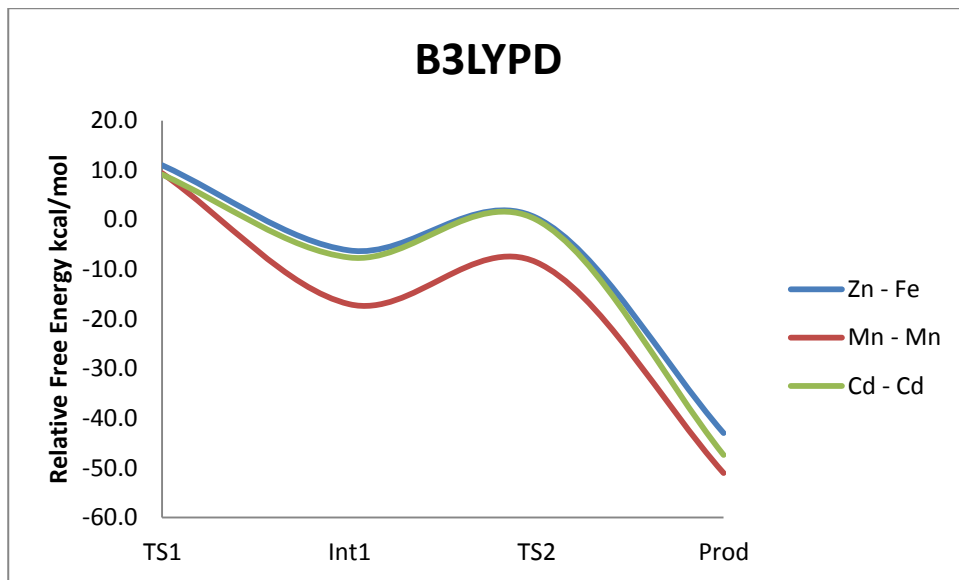


Figure 6- Potential energy surface obtained for the three models with the B3LYPD density functional and 6-311++G (2d,2p) basis-set. Values for the phosphotriesterase reaction mechanism.

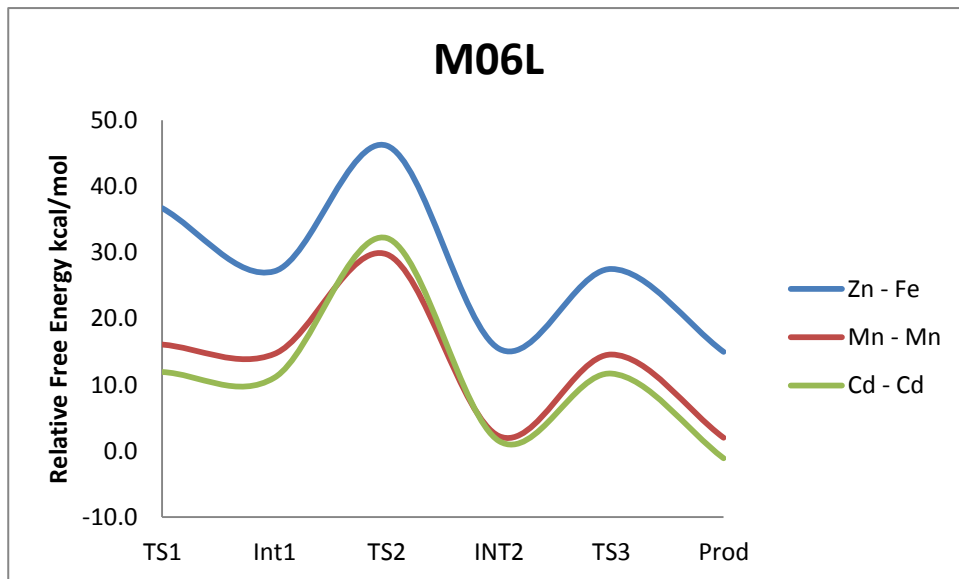


Figure 7- Potential energy surface obtained for the three models with the M06L density functional and 6-311++G (2d,2p) basis-set. Values for the phosphodiesterase reaction mechanism.

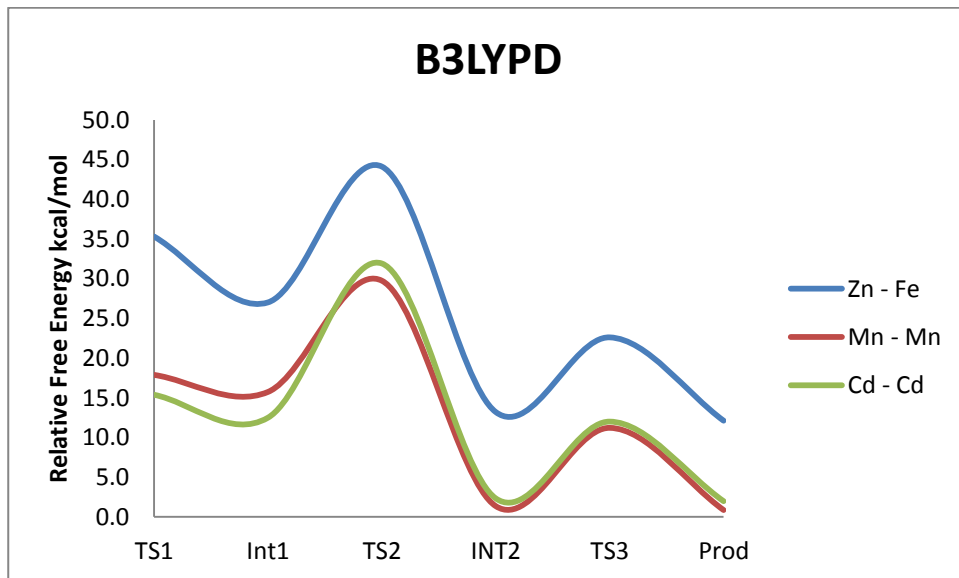


Figure 9- Potential energy surface obtained for the three models with the B3LYPD density functional and 6-311++G (2d,2p) basis-set. Values for the phosphodiesterase reaction mechanism.

			ZnFe	Mn	Cd
			Bond distances Å		
ES	O99	Metal45	2.405	2.320	2.400
	P96	O98	1.610	1.610	1.607
	O118	Metal44	1.942	1.951	2.202
	O118	Metal45	2.104	2.065	2.218
	O118	P96	3.562	3.654	4.026
	O98	H121	1.986	1.981	1.918
	O118	H119	0.968	0.969	0.973
	H119	O120	3.938	3.661	2.022
	O120	H121	0.974	0.974	0.975
	Metal44	O49	4.794	4.742	5.427
Metal45	O49	2.080	2.053	2.336	
TS1	O99	Metal45	2.055	1.998	2.330
	P96	O98	1.617	1.629	1.620
	O118	Metal44	1.988	2.028	2.233
	O118	Metal45	2.724	2.896	2.323
	O118	P96	1.948	1.887	2.095
	O98	H121	3.191	3.146	3.053
	O118	H119	0.973	0.978	0.973
	H119	O120	2.297	2.008	2.153
	O120	H121	0.978	0.980	0.978
	Metal44	O49	4.121	4.200	4.509
Metal45	O49	2.027	2.079	2.294	
Int	O99	Metal45	2.057	1.994	2.192
	P96	O98	1.759	1.737	1.754
	O118	Metal44	2.217	2.303	2.506
	O118	Metal45	3.534	3.771	3.652
	O118	P96	1.678	1.670	1.672
	O98	H121	1.900	1.870	1.910
	O118	H119	1.006	1.003	1.000

	H119	O120	1.703	1.750	1.742
	O120	H121	0.980	0.979	0.980
	Metal44	O49	1.968	2.016	2.190
	Metal45	O49	2.072	2.001	2.223
TS2	O99	Metal45	2.063	1.994	2.267
	P96	O98	1.885	1.866	1.885
	O118	Metal44	2.157	2.198	2.403
	O118	Metal45	3.459	3.700	3.327
	O118	P96	1.600	1.599	1.593
	O98	H121	1.410	1.407	1.455
	O118	H119	1.321	1.318	1.401
	H119	O120	1.139	1.143	1.089
	O120	H121	1.079	1.074	1.064
	Metal44	O49	1.969	1.999	2.193
	Metal45	O49	2.075	1.995	2.263
Prod	O99	Metal45	2.150	2.110	2.316
	P96	O98	5.512	5.399	5.395
	O118	Metal44	2.041	2.072	2.386
	O118	Metal45	3.491	3.460	3.202
	O118	P96	1.505	1.508	1.519
	O98	H121	0.986	0.986	0.986
	O118	H119	2.633	2.545	1.984
	H119	O120	0.976	0.974	0.979
	O120	H121	1.821	1.820	1.826
	Metal44	O49	2.027	2.077	2.240
	Metal45	O49	2.056	1.984	2.238

Table 1- Important distances for the first reaction mechanism. Three models presented.

			Mn	Cd	ZnFe
			Distances Å		
ES	O99	Metal45	2.101	2.224	2.065
	O98	Metal44	2.081	2.361	1.984
	O49	H113	1.659	1.752	1.649
	O50	H79	1.794	1.937	1.798
	O97	H95	4.443	2.417	4.924

	O50	H82	2.755	2.432	2.848
	H95	O50	0.970	0.974	0.970
	O97	H82	3.505	2.944	3.466
	Metal44	O49	3.476	4.909	3.992
	Metal45	O49	2.051	3.278	2.001
	N31	H82	1.046	1.025	1.050
	N28	H79	1.035	1.027	1.035
	O50	H79	1.794	1.937	1.798
	O49	P96	3.428	3.905	3.503
	O97	P96	1.613	1.624	1.614
TS1	O99	Metal45	2.079	2.203	1.994
	O98	Metal44	2.058	2.218	1.940
	O49	H113	1.638	1.638	1.890
	O50	H79	1.822	1.822	1.823
	O97	H95	2.755	2.756	2.427
	O50	H82	2.338	2.385	2.362
	H95	O50	0.970	0.970	0.973
	O97	H82	4.056	2.081	2.070
	Metal44	O49	3.015	2.895	2.146
	Metal45	O49	2.144	2.259	2.238
	N31	H82	1.108	1.019	1.019
	N28	H79	1.031	1.013	1.028
	O50	H79	1.822	1.822	1.823
	O49	P96	2.001	2.001	2.001
	O97	P96	1.670	1.676	1.665
Int1	O99	Metal45	2.115	2.237	2.051
	O98	Metal44	2.022	2.171	1.946
	O49	H113	3.582	3.478	3.680
	O50	H79	2.042	2.033	2.030
	O97	H95	3.476	3.579	3.411
	O50	H82	1.974	1.988	1.986
	H95	O50	0.969	0.969	0.969
	O97	H82	3.069	2.539	2.468
	Metal44	O49	3.464	3.549	3.433
	Metal45	O49	4.021	4.010	4.019
	N31	H82	1.021	1.021	1.020

	N28	H79	1.020	1.021	1.020
	O50	H79	2.041	2.033	2.030
	O49	P96	1.643	1.647	1.627
	O97	P96	1.724	1.733	1.726
TS2	O99	Metal45	2.116	2.230	2.073
	O98	Metal44	2.053	2.196	2.010
	O49	H113	3.280	3.157	3.126
	O50	H79	1.160	1.167	1.155
	O97	H95	1.233	1.238	1.273
	O50	H82	2.870	2.869	2.943
	H95	O50	1.207	1.203	1.169
	O97	H82	2.804	2.772	2.694
	Metal44	O49	3.478	3.558	3.509
	Metal45	O49	4.001	4.066	3.835
	N31	H82	1.018	1.020	1.022
	N28	H79	1.354	1.350	1.360
	O50	H79	1.160	1.166	1.155
	O49	P96	1.621	1.623	1.622
	O97	P96	1.932	1.918	1.918
Int2	O99	Metal45	2.225	2.338	1.986
	O98	Metal44	2.116	2.254	2.004
	O49	H113	3.680	3.653	3.361
	O50	H79	1.006	1.720	1.008
	O97	H95	0.991	0.992	0.991
	O50	H82	3.044	3.052	3038.000
	H95	O50	1.721	1.720	1.725
	O97	H82	1.883	1.881	1.890
	Metal44	O49	3.813	3.918	3.886
	Metal45	O49	4.178	4.224	4.398
	N31	H82	1.028	1.028	1.027
	N28	H79	1.781	1.788	1.768
	O50	H79	1.006	1.005	1.008
	O49	P96	1.573	1.578	1.571
	O97	P96	6.224	6.283	5.991
TS3	O99	Metal45	2.226	2.340	1.979

	O98	Metal44	2.097	2.238	1.999
	O49	H113	3.801	3.752	3.579
	O50	H79	1.549	1.551	1.584
	O97	H95	1.208	1.205	1.216
	O50	H82	2.801	2.805	2.807
	H95	O50	1.209	1.212	1.201
	O97	H82	1.439	1.444	1.426
	Metal44	O49	3.816	3.930	3.887
	Metal45	O49	4.158	4.198	4.386
	N31	H82	1.122	1.119	1.128
	N28	H79	1.083	1.083	1.074
	O50	H79	1.549	1.551	1.584
	O49	P96	1.577	1.583	1.575
	O97	P96	6.148	6.171	5.994
Prod	O99	Metal45	2.220	2.337	1.984
	O98	Metal44	2.110	2.246	2.002
	O49	H113	3.609	3.459	3.392
	O50	H79	1.864	1.875	1.893
	O97	H95	1.171	1.713	1.698
	O50	H82	2.831	2.836	2.833
	H95	O50	0.996	0.995	0.997
	O97	H82	1.006	1.006	1.008
	Metal44	O49	3.829	3.937	2.888
	Metal45	O49	4.156	4.196	4.389
	N31	H82	1.741	1.740	1.789
	N28	H79	1.027	1.026	1.025
	O50	H79	1.864	1.875	1.893
	O49	P96	1.574	1.579	1.572
	O97	P96	6.126	6.132	5.889

Table 2- Important distances for the second reaction mechanism. Three models presented.

6 Scope - Tyrosine Ammonia-Lyase given by a QM/MM model

Tyrosine Ammonia Lyase (TAL) catalyzes the deamination of tyrosine to p-coumaric acid in purple phototropic bacteria and Actinomycetales. The enzyme is used in bioengineering and has potential to be used industrially. It belongs to a family of enzymes that uses a 4- methylidene-imidazole-5-one (MIO) cofactor to catalyze the deamination amino acids. In the present work, we used a 5216 QM/MM model of TAL to explore two putative reaction paths for its catalytic mechanism. The N-MIO mechanism was previously discussed in the literature and part of it was already been studied computationally. We improved on previous studies by using a larger, more complete model of the enzyme, and by describing the complete reaction path. The activation energy for this mechanism, in agreement with the previous study, is 28.5 kcal/mol. We also found another reaction path that has overall better kinetics and reaches the products in a single reaction step. The barrier for this Single-Step mechanism is 16.6 kcal/mol, which agrees very well with the experimental k_{cat} of 15.6 kcal/mol.

This work has been published in the Archives of Biochemistry and Biophysics journal.

7 Scope – Human pancreatic α -amylase

The human pancreatic α -amylase is an enzyme that catalyzes the cleavage of glycosidic bonds through hydrolysis. It is part of the Glycoside Hydrolase (GH) family of enzymes. The enzymes on this family perform a wide variety of reactions involving glycosidic bonds from glycosylation and deglycosylation to hydrolysis and cyclisation. There was a proposed mechanism for this enzyme but no computational studies were made before to give us full atomistic insight of the mechanism presented in this work. In the present work we tested different protonation states of the amino acid responsible for initiating the catalytic mechanism and molecular dynamics simulations were carried out to ensure the correct initial conformation before proceeding to the QM/MM calculations.

In the following study we present and discuss the obtained results and the two mechanistic steps. At the end we conclude the study by proposing a catalytic mechanism with a small difference of the previously proposed mechanism.

Further information about the study will be found in the next pages of this document.

This work has been published on JCPC.

8 Scope - Organo-Phosphate Degrading Enzyme

Phosphoesterases have been long known to catalyze the hydrolysis of a number of synthetic phosphotriesters normally used as insecticides. In our work we report that the promiscuity of on phosphoesterase (Organo-Phosphate Degrading Enzyme from *Agrobacterium radiobacter* P230) goes even further. Other than proposing a catalytic mechanism for its catalytic power as a Phosphotriesterase, we propose a phosphodiesterase catalytic mechanism for the same enzyme.

All calculations were performed using Density Functional Theory (DFT) choosing for optimization in gas phase the B3LYP functional, The 6-31G(d,p) basis set was used to describe heavy elements and hydrogen atoms and the LANL2DZ pseudopotential transition metal ions.

We performed computations with an active site model of 122 atoms built from the x-ray structures. The results obtained show that the catalysis of a phosphotriesters, a two-step mechanism, differs from that of the phosphodiester, a three step mechanism.

Antiferromagnetic singlet high spin configuration is practically isoenergetic to that of ground ferromagnetic septet spin state.

Further information about this enzymatic mechanism may be found in the following pages.

This work has been published in CHEMISTRY – An European Journal.



New insights in the catalytic mechanism of tyrosine ammonia-lyase given by QM/MM and QM cluster models



Gaspar P. Pinto^{a,b,1}, António J.M. Ribeiro^{a,b,1}, Maria J. Ramos^a, Pedro A. Fernandes^a, Marirosa Toscano^b, Nino Russo^{b,*}

^a UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal

^b Dipartimento di Chimica, Università della Calabria, 87036 Arcavacata di Rende, Italy

ARTICLE INFO

Article history:

Received 19 December 2014
and in revised form 16 February 2015
Available online 12 March 2015

Keywords:

Tyrosine ammonia lyase
QM
QM/MM
Catalytic mechanism
DFT
Enzymatic catalysis

ABSTRACT

Tyrosine ammonia lyase (TAL) catalyzes the deamination of tyrosine to p-coumaric acid in purple phototropic bacteria and Actinomycetales. The enzyme is used in bioengineering and has the potential to be used industrially. It belongs to a family of enzymes that uses a 4-methylidene-imidazole-5-one (MIO) cofactor to catalyze the deamination amino acids. In the present work, we used a QM/MM and a QM cluster models of TAL to explore two putative reaction paths for its catalytic mechanism. Part of the N-MIO mechanism was previously studied by computational methods. We improved on previous studies by using a larger, more complete model of the enzyme, and by describing the complete reaction path. The activation energy for this mechanism, in agreement with the previous study, is 28.5 kcal/mol. We also found another reaction path that has overall better kinetics and reaches the products in a single reaction step. The barrier for this Single-Step mechanism is 16.6 kcal/mol, which agrees very well with the experimental k_{cat} of 16.0 kcal/mol. The geometrical parameters obtained for the cluster and QM/MM models are very similar, despite differences in the relative energies. This means that both approaches are capable of describing the correct catalytic path of TAL.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Tyrosine ammonia lyase (TAL) was first discovered in 2001 from the genome analysis of *Rhodobacter capsulatus*. It catalyzes the deamination of tyrosine to p-coumaric acid with a turnover number of 27.7 s^{-1} [1]. The reaction catalyzed by TAL is of biological, bioengineering, and industrial interest: p-coumaric acid is the chromophore of photoactive yellow protein of purple phototropic bacteria [2] and a precursor in the biosynthesis of caffeic acid in Actinomycetales; [3] the enzyme is also used in bioengineered pathways to produce flavonoids [4] and resveratrol [5]; finally, TAL has the potential to be used in an industrial setting in the production of several industrial chemicals whose precursor is p-coumaric acid [6,7].

Tyrosine ammonia lyase together with phenylalanine ammonia lyase (PAL) and histidine ammonia lyase (HAL) form a small group of enzymes that have a similar active center and catalyze the deamination of amino acids. These enzymes have a common electrophilic cofactor denominated 4-methylidene-imidazole-5-one

(MIO) [8] which is formed from the autocatalytic cyclization of three adjacent amino acids, Ala149-Ser150-Gly151 in TAL [9], during the folding of the enzymes [10]. PAL and TAL, specifically, are so similar that a single amino acid, His89, has been identified as selectivity switch between them [11]. In TAL, His89 is making an hydrogen bond with the hydroxyl group of the tyrosine substrate, by mutating this residue to a non-polar (phenylalanine) or more bulky (glutamine) residue, TAL loses the catalytic power over tyrosine but gains it over phenylalanine. The opposite mutation on a PAL makes it capable of catalyzing the deamination of tyrosine. Apart from the MIO cofactor, Tyr60 and Tyr300 have also been identified by mutagenesis studies as essential for catalysis in TAL [12]. In HAL, Tyr53 and Tyr280 are equivalent to the two TAL tyrosine amino acids [8], and in PAL, Tyr371 is in the same position as the Tyr300 of TAL [13]. From these resemblances, it is expected that the reaction mechanism followed by the three enzymes is similar.

Initially, two mechanisms were proposed for the deamination of amino acids by MIO enzymes: a Friedel–Craft type mechanism where the MIO cofactor (thought to be dehydroalanine at the time) forms a covalent intermediate with the aromatic ring of the substrate [14,15]; and a mechanism where the cofactor does a covalent intermediate with the amine group of the substrate (N-MIO) [16]. Recently, a computational study of the TAL reaction

* Corresponding author.

E-mail address: nrusso@unical.it (N. Russo).

¹ António J.M. Ribeiro and Gaspar P. Pinto contributed equally to this work.

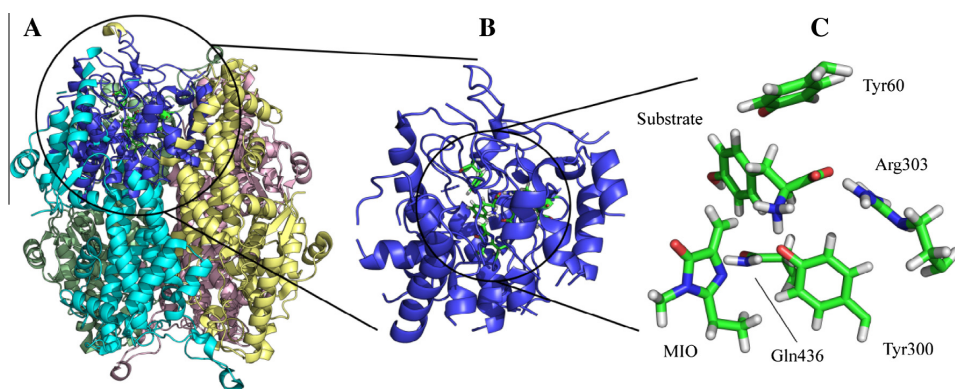


Fig. 1. A homotetramer of the tyrosine ammonia lyase enzyme, and the model used in this work. (A) – Homotetramer of TAL colored by chain. The dark blue portion, is the part of the enzyme we used in the ONIOM calculations. (B) – QM/MM model used in this work. In blue ribbons is represent the MM layer of the model, and in sticks is represent the QM layer. (C) – Detailed representation of the QM layer used in the ONIOM calculations.

mechanism showed that the N-MIO mechanism is much more favorable than the Friedel–Crafts mechanism [17]. In that paper, the authors describe two steps of the N-MIO reaction path.

In the present work we used two models of the TAL enzyme to explore its potential energy surface in great detail: a 172 atoms DFT cluster model and a 5216 atoms QM/MM model (see Fig. 1). We were able to describe, with both models, the complete reaction of the N-MIO mechanism, from the state of having a protonated-amine substrate from the leaving of a protonated amine group. Furthermore, we found an additional mechanism that is compatible with the TAL active center and has better kinetics than the N-MIO mechanism. In this new mechanism, there is no covalent intermediate with the cofactor and the reaction happens in a single reaction step with a barrier of 16.6 kcal/mol.

Methods

The QM/MM model of TAL was built from the 2O7B X-ray structure [13]. This structure has the entire TAL homotetramer, and a *p*-coumarate molecule bonded to the active center. For the QM/MM model (see Fig. 1), we started by selecting all the residues within 20 Å of the substrate. In order to minimize the number of cuts and make the model more cohesive, some residues were added to this selection to fill small gaps in the selection (residues missing in the middle of an alpha helix, for example). The water molecules picked up by the selection where left in the model. The tyrosine substrate was modeled from the *p*-coumarate molecule by adding the amine group to the α -carbon, with attention made to the ensuing chirality (S at the α -carbon, a *L*-tyrosine). This model of the enzyme has 5216 atoms. To establish the QM/MM division, residues deemed more important to catalysis where included in the quantum mechanics (QM) layer: the tyrosine substrate, tyrosine 60, tyrosine 300, arginine 303, the MIO cofactor, and glutamine 436; for a total of 99 atoms. The rest of the atoms where included in the molecular mechanics (MM) layer. The selection of the 20 Å sphere around the substrate was done on PYMOL, [18] while the modeling task and the division of the QM/MM system was done on GAUSSVIEW [19].

We used the antechamber program, [20] part of the AMBER10 suite of programs, [21] to parameterize the unbound tyrosine substrate and the MIO cofactor, since the amber force field does not include parameters for these molecules. Antechamber attributes parameters from the GAFF force field [22] to bonds, angles, dihedrals, and van der Waals radii. Atomic charges were calculated from a QM calculation at the HF/6-31G(d) level, [23] to maintain

compatibility with the amber force field, followed by a RESP (Restrained ElectroStatic Potential) calculation [24].

The cluster model was built from the optimized structure of the reactants of the QM/MM model in order to ensure the similarity of the initial geometries. The side chains of the most important residues for the catalytic mechanism were maintained: the tyrosine substrate, tyrosine 60, tyrosine 300, arginine 303, the MIO cofactor, phenylalanine 66, histidine 86, phenylalanine 350, methionine 405 (chain B), asparagine 432 and glutamine 436. The modeling task was done on GAUSSVIEW [19].

The QM layer of the QM/MM model and the QM cluster model were treated with the B3LYP [25,26] density functional and the 6-31G(d) basis-set [27]. The MM layer of the QM/MM model was treated with the AMBER03 force field [28] as implemented in GAUSSIAN09. All calculations were done with the GAUSSIAN09 software [29].

The approach for exploring the potential energy path was the same for both models. Structures of transition states and intermediates were first obtained by scanning appropriate coordinates along the reaction path. Afterwards, all states, including minima and transition states, were freely optimized. During optimizations, the only constrained atoms in the QM/MM model belong to low layer residues where the cuts were made in the preparation of the model. The coordinates of the atoms of these residues were frozen in place during all optimizations. In the cluster model, the atoms of the residue side chains where the cuts were made were also frozen in space to ensure that the structure of the model continued the same throughout the calculations, to simulate the effect of the protein scaffold. Frequency calculations were done to confirm the position (as minima or maxima) of all obtained states. In the QM/MM model, zero point energy and entropic corrections were added to the potential energy path, in order to obtain an approximated free energy path. In order to improve the energy values, single point energy calculations were done for both models at the B3LYP/6-311++G(2d,2p) level [30–32]. Since explicit water molecules near the active center were present in the QM/MM model, and the bulk solvent was at least at 20 Å from of the active center, no additional solvent corrections were required.

Results and discussion

From the early exploration of the reaction space associated with the active center of tyrosine ammonia lyase we found two potential reaction paths. The first one, presented in Scheme 1, had been

previously proposed in the literature [16]. It involves the attack of the amine group of the substrate to the cofactor. While we were developing this work, another group published a paper concerning this same mechanism [17]. We compare and discuss our results against this paper on the proper section. The second mechanism we found, shown in Scheme 2, has not been proposed previously. It is a simple mechanism where the MIO cofactor is not involved directly. The energies of this alternative mechanism, however, are more reasonable than that of the first, as will be discussed. In agreement with the previous theoretical study on TAL, our models are also not compatible with the Friedel–Crafts mechanism, proposed initially for the TAL reaction.

The N-MIO mechanism

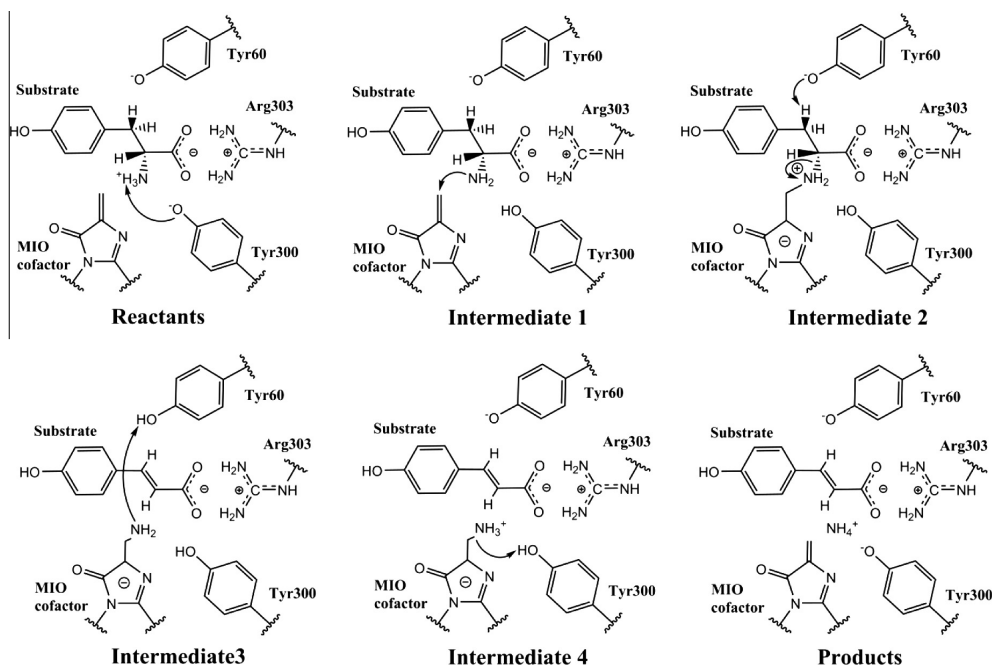
The N-MIO mechanism occurs through a reaction path with five transition states. The obtained structures and energies for this mechanism are shown in Figs. 2 and 3, respectively. In the reactants state, the nitrogen of the amine group is at 4.228|3.063 Å (cluster model distance|QM/MM model distance) of the methylene carbon of the MIO cofactor, and the hydroxyl oxygen of Tyr60 is at 2.436|2.027 Å of the proton of the substrate to be removed in a later step. The hydroxyl oxygen of Tyr300 is interacting very strongly with the amine group, at 1.683|1.454 Å. The two catalytic tyrosine residues, Tyr60 and Tyr 300, are deprotonated at this point, while the amine group of the substrate is protonated. It is not clear, nor has been explained by anyone that published previously about this mechanism, how is it possible to have two tyrosine amino acids simultaneously deprotonated in the same active center without any strong charge compensation. Upon substrate binding, Tyr300 is stabilized by the protonated amine group of the substrate, and so, it is possible that this residue only loses its proton (to the surrounding waters, for example) after substrate binding. The doubt remains for Tyr60, though.

The first step of the mechanism is, precisely, the transference of a proton from the amine group to the hydroxyl oxygen of Tyr300. At the transition state (TS1 in Fig. 2) the proton is in the middle ground between Tyr300 (1.276|1.369 Å) and the amine (1.239|1.163 Å). The activation energy of this step, at the B3LYP/6-311++G(2d,2p) level, is 6.62.5 kcal/mol.

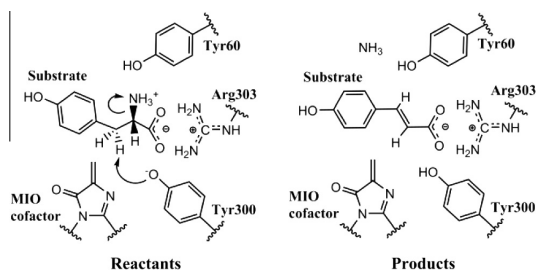
In the intermediate after TS1 (INT1 in Fig. 2), Tyr300 is protonated and the amine group is neutral. The transferred proton is still making a close hydrogen bond (1.683|1.587 Å) to the amine. The distances of the cofactor to the amine are the same as they were in the reactants and the Tyr60 distance to the substrate proton has elongated slightly, pulled by one amine proton. This state is slightly more stable than the reactants state (−4.7|−0.6 kcal mol^{−1}).

The second reaction step (TS2 in Fig. 2) is the nucleophilic attack of the amine group of the substrate to the cofactor. In this transition state, the nitrogen of the amine group and the methylene carbon of the cofactor are at a distance of 2.018|2.432 Å. The proton of the substrate is still bonded and interacting with the side chain of Tyr60 at 2.212|2.294 Å. The activation energy for this step (in relation to INT1) is 9.7|8.8 kcal/mol and 5.0|8.2 kcal/mol in relation to the reactants state.

At the second intermediate (INT2 in Fig. 2) there is a covalent bond between the MIO cofactor and the substrate mediated by the amine group. At this point, the hydrogen of the phenol group of Tyr300 changes its orientation and makes a hydrogen bond with a nitrogen atom of the cofactor. This interaction is meaningful, since in this state the cofactor is negatively charged and there are no positive residues around it to stabilize that charge. This intermediate is very stable in the QM/MM model, with an energy of 21.0 kcal/mol in relation to the previous intermediate, and −21.6 kcal/mol, in relation to the reactants. The QM model does not show a stabilization of the same magnitude. This intermediate has an energy difference of −0.8 kcal/mol in relation to the previous intermediate and −5.5 kcal/mol in relation to the reactants.



Scheme 1. Schematic representation of the N-MIO mechanism reaction path.



Scheme 2. Schematic representation of the Single-Step mechanism reaction path.

On the third transition state (TS3 in Fig. 3), the amine group breaks its bond from the substrate, while Tyr60 accepts a proton from the substrate. This is the main step in the reaction and is also the rate limiting step. The proton of the substrate that is transferred to Tyr60 is already completely bonded to it, but still very close to the substrate, at 1.896|2.028 Å. The bond between the amine and the substrate is elongated but not broken (1.791|2.051 Å). The hydrogen bond between Tyr300 and the cofactor remains in place. This barrier is 34.6|28.4 kcal/mol high in relation to the previous intermediate, and 29.1|6.8 kcal/mol in relation to the initial reactants. It is important to distinguish between these two energies, because to calculate the k_{cat} we must consider the lower state, corresponding to intermediate 2, and not the initial energy, corresponding to the reactants: the limiting step is defined by the path highest barrier. The turnover number is calculated in the proper section, below.

From transition state 3, the path goes to intermediate 3.1 (INT 3.1 in Fig. 2). This minimum was only observed in the QM/MM model calculations.

We found that for the next reaction step to occur in the QM/MM model, the proton of Tyr60 must be near the amine group of the

cofactor. In order for this to happen, the side chain of Tyr60 has to move significantly. We induced that movement by scanning the distance between the terminal proton of Tyr60 and the nitrogen of the amine group. After a slight rearrangement of the structure a new energy minima is achieved (INT3.2 in Fig. 2). In relation to the intermediate 3.1, this one is 2.0 kcal/mol more stable. The cluster model intermediate 3 corresponds to INT 3.2 of the QM/MM model. At this point, the p-coumarate product is already formed, but the amine group remains covalently bonded to the enzyme by the MIO cofactor. Tyr60 is completely bonded to the proton abstracted from the substrate. That proton is now at a distance of 1.778|1.774 Å from the amine group. Tyr300 remains tightly interacting with the cofactor through a hydrogen bond. The energy of this intermediate is $-10.5|-3.2$ kcal/mol in relation to intermediate 2 and it is $-16.0|-18.4$ kcal/mol more stable than the reactants.

The next transition state (TS4 in Fig. 2) is a proton transfer between Tyr60 and the amine group bonded to the cofactor. At this point the proton is at 1.243|1.226 Å of the phenol oxygen of Tyr60 and at 1.339|1.330 Å of the nitrogen atom of the amine group. As usual for a proton transfer this step amounts to a small barrier of 4.8|2.4 kcal/mol in relation to the previous step. In relation to the initial reactants, this step has $-11.2|-16.0$ kcal/mol. It is unclear from our results if in this and the next step the p-coumarate product is still present in the active center. In principle both, TS4 and TS5 could occur without any help of p-coumarate. In the specific case of TS5, p-coumarate may be more of an impediment than an aid, as is explained below.

The next minimum in the reaction path (INT4 in Fig. 2) is characterized by a protonated amine group bonded to the cofactor. Now, the cofactor is no longer negatively charged, although it maintains a strong dipole between the amine group, with a formal charge of +1, and the ring portion with a formal charge of -1. For this reason, the hydrogen bond between Tyr300 and the cofactor is no longer stable, and Tyr300 prefers to make a hydrogen bond with

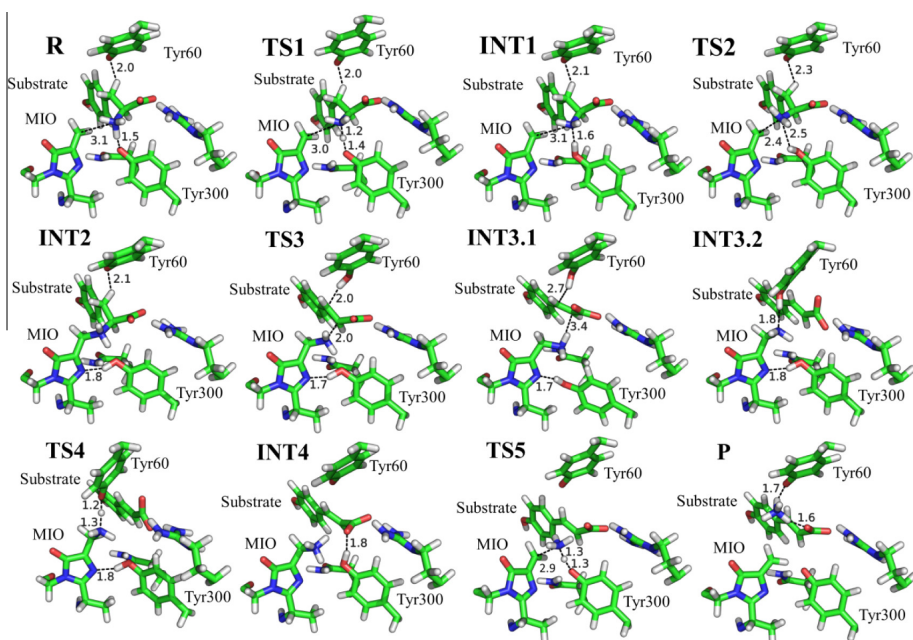


Fig. 2. Reaction path for the N-MIO mechanism as studied by the QM/MM model. The represented atoms are the ones in the QM layer of the QM/MM model.

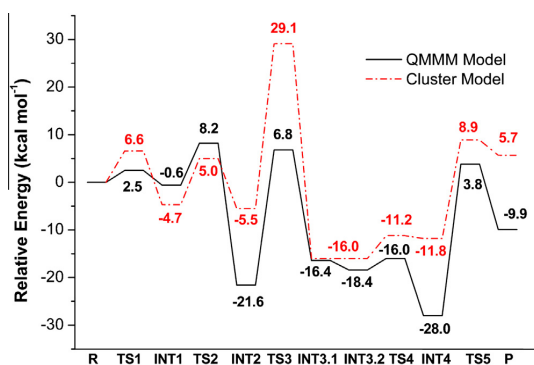


Fig. 3. Energy profile for the N-MIO mechanism, as calculated with the QM/MM (free-energy) and QM models (potential energy).

the carboxylic group of the substrate. The hydrogen bond between the hydroxyl oxygen of Tyr300 and the protonated amine is maintained, nevertheless. The behavior of Tyr60 in the two models is distinct in this intermediate. In the QM/MM model, Tyr60 returns to its initial position, after donating its proton. This intermediate is extremely stable, with an energy of -9.6 kcal/mol in relation to the previous intermediate and -28.0 kcal/mol in relation to the reactants. In the cluster model, however, Tyr60 stays close to amine group. This prevents the intermediate to have a big stabilization in relation to the previous transition state. The difference in energy for the cluster model is 0.6 kcal/mol.

In the last transition state (TS5 in Fig. 2) the amine group detaches from the cofactor while taking a proton from Tyr300: the methylene carbon of the cofactor is at $3.264|2.871$ Å of the nitrogen of the amine group. The transferred proton is at $1.281|1.252$ Å of the nitrogen of the amine group and at $1.231|1.253$ Å of the phenol oxygen of Tyr300. The activation energy of this step is $20.7|31.8$ kcal/mol. The QM/MM barrier is very high, and not compatible with the general knowledge of enzymatic catalysis. We hypothesize that, if this is in fact the reaction path followed by the enzyme, the presence of the substrate may be hindering the movement of the amine group away from the cofactor. It is entirely possible that at this point the deaminated product has already left the active center, and that the abstraction of the amine group from the cofactor takes place in the hollow active center. For this reason, we do not consider this step the limiting step of the reaction, although it corresponds to the highest energy barrier. We did not try to model the reaction in these conditions because it is expected that after the departure of the products, the active center becomes filled with water molecules. Since such structure is not available, and it is difficult to model such heterogeneous arrangement of molecules, we opt to use the model we already have, with the product in the active center.

In the final structure (P in Fig. 2) both the ammonium ion and p-coumarate are free and in their final protonation state. The protein is also completely regenerated, with a structure and protonation states equal to the reactants state: with the two tyrosine amino acids deprotonated and a deaminated cofactor. The positively charged ammonium ion is being stabilized by two hydrogen bonds with the negatively charged Tyr60 ($1.626|1.659$ Å) and the carboxylic group of the product ($1.587|1.580$ Å). This state is more energetic than the previous intermediate by $17.5|18.1$ kcal/mol and has $5.7|$ – 9.9 kcal/mol in relation to the reactants.

The Single-Step mechanism

While exploring the conformational space of the TAL active center, we come across another possible mechanism that was not previously encountered in the literature. This alternative mechanism, which we named Single-Step mechanism is simpler and has more favorable kinetic and thermodynamic energies. However, the role for the cofactor is not as clear as in the previous mechanism. This mechanism is shown in Scheme 2 and Fig. 4, and the calculated energies in Fig. 5.

In the reactants state of this model (R in Fig. 4) the substrate is in a slightly different conformation than previously: instead of facing the cofactor, the amine group of the substrate is facing Tyr60. Since it is a planar p-coumarate molecule that is present in the active center of TAL in the crystallographic structure, both positions for the amine are possible. We tried different protonation states for both tyrosine residues and found that the only one that leads to sensible kinetics is the state where Tyr60 is protonated and Tyr300 deprotonated. In the QM/MM model, Tyr60 is interacting with the amine group by two loosely hydrogen bonds of 2.523 and 2.754 Å, while Tyr300 is at a distance of 2.758 Å of the proton to be removed from the substrate. In the QM model Tyr60 is at a distance of 2.841 Å of one proton from the amine group, but Tyr300 is already very close (1.893 Å) to the proton that is going to be removed from the substrate. The only possible role for the cofactor in this mechanism is the stabilization of the negatively charged Tyr300, with the help of Gln436. In relation to the previous mechanism reactants state, this one seems more probable, since only one of the tyrosine amino acids is deprotonated.

In the transition state (TS in Fig. 4) the proton of the substrate is in the middle of the substrate and Tyr300. The proton is at $1.555|1.523$ Å of the carbon of the substrate it was bonded to, and at $1.173|1.164$ Å of the phenol oxygen of Tyr300. The bond between the substrate and the amine group is slightly elongated, $1.635|1.564$ Å from the initial $1.532|1.512$ Å, but not yet broken. Tyr60 is still doing the two hydrogen bonds with the amine group at 2.507 and 2.494 Å each in the high layer of the QM/MM model, and one hydrogen bond (2.751 Å) in the QM model. The activation energy for this step is $18.1|16.6$ kcal/mol. Since this reaction path has only one transition step, this is the rate limiting step.

In the third state (P in Fig. 4) the products are already formed. The amine group leaves as a neutral ammonia molecule. The ammonia may be protonated when it leaves the active center or still in it, since there are a lot of water molecules in the vicinity. In fact, in this state, the ammonia is making a hydrogen bond with a water molecule in the low layer. Like in the previous mechanism, Tyr300 is making a close hydrogen bond ($1.471|1.629$ Å) with the carboxylic group of p-coumarate. To recover its original state, Tyr300 needs to be deprotonated. This may be achieved by losing its proton to a water molecule, or even to the ammonia, after p-coumarate had left the active center. The energy of this step relative to the reactants is $-32.2|$ – 43.1 kcal/mol. This energy is very low, but should not be interpreted as a purely chemical energy i.e. the difference does not come purely from the transformation of the reactants in products but also from rearrangements in the enzyme structure. The ammonia molecule, particularly, goes to a completely different position after detaching from the substrate.

Comparison of the QM/MM model and the cluster QM model

The comparison between the two computational models used in this work provides interesting insights about the strengths and limitations of both approaches. The cluster model is easier to prepare, although greater thought must be put on the choice of the residues to include, since the ones that are left out will be

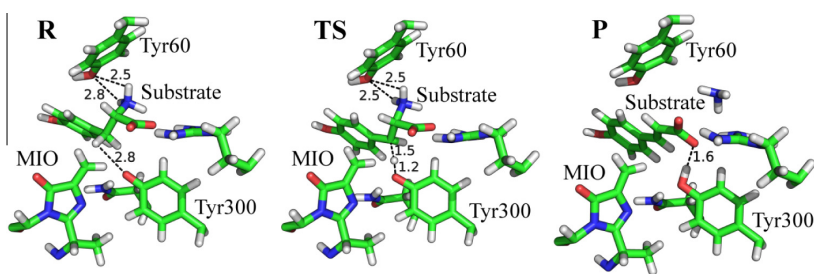


Fig. 4. Reaction path for the Single-Step mechanism as studied by the QM/MM model. The represented atoms are the ones in the QM layer in the ONIOM calculations.

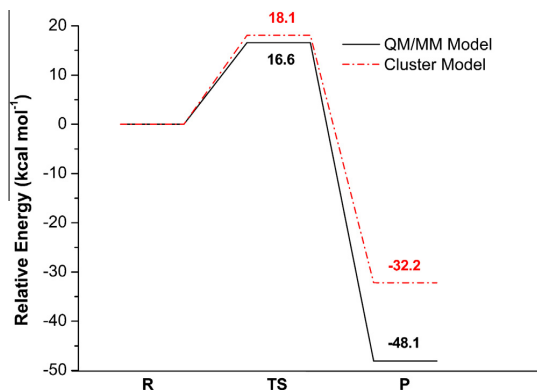


Fig. 5. Energy profile for the Single-Step mechanism as calculated by the QM/MM (free-energy) and the cluster (potential energy) models.

completely neglected. The QM/MM model requires parameters for all the atoms in the MM layer. All residues not included in the base force field must be parametrized, which is an additional laborious step. Furthermore, the complexity and empirical character of molecular mechanics parameters can be a source of errors or imprecisions in the model.

The distances and angles of the geometries obtained for the reaction paths of both models are very similar. The energetic profiles of the models are also similar, overall, but some stationary states have significant differences. We will try to understand such deviations in this section.

In the N-MIO reaction mechanism, both models start with the same positioning of Tyr60 and Tyr300 in relation to the substrate. Tyr300, the most active catalytic residue in this step, is interacting closely with the substrate through the amine's proton that will be removed (1.683 Å|1.454 Å) (cluster model distance|QM/MM model distance). Tyr60 is not going to participate actively in the first step but it is already interacting with the substrate through a H-bond (2.436 Å|2.027 Å) with the hydrogen that will be removed from the substrate in a later step. The distance between the amine of the substrate and the nearest carbon atom of the cofactor is distinct, (4.228 Å|3.063 Å), however, this distance is not relevant to the first step of the mechanism.

At the transition state, both the bond that is forming (1.239 Å|1.163 Å) and the bond that is breaking (1.276 Å|1.369 Å) have equivalent lengths in both models. In INT1, Tyr300 interacts very tightly with the amine (1.683 Å|1.587 Å) without distinction. The relative energies of the transition state and intermediate of this step are mostly equivalent and in accordance with typical

proton transfers, although the values from the cluster model are more spread. The transition state energies are 6.6 kcal/mol for the cluster model and 2.5 kcal/mol for the QM/MM. At INT1, the cluster model yielded a lower value -4.7 kcal/mol in relation to the QM/MM model, -0.6 kcal/mol. These differences can be explained by the stabilization provided to the charged substrate and Tyr300 in the reactants and transition states. Due to the surrounding charges of the enzyme scaffold, the QM/MM model is able to stabilize the active center of these two states better than the cluster model, even with the implicit solvent taken into account. That is why the neutrally charged INT1 (in relation to the amine of the substrate and Tyr300) is more stable in the cluster model, in relation to the other two states.

The second step, which starts with INT1, presents equivalent activation energies for both theories: 9.7 kcal/mol for the cluster model and 8.8 for the QM/MM model. On the INT2 state, however, the QM/MM model is 16.1 kcal/mol more stable than the cluster model. We do not detect any significant differences in the geometries of the active center. Once more, since this is a charged state, we think that the QM/MM model is able to stabilize it much better than the cluster model. More specifically, we found that a chain of hydrogen bonds extends from the positively charged amine group to an asparagine of the low layer, through a linking water molecule.

The energetic cost of TS3 is high in both models, in spite of the high energy differences in INT2:

34.6 kcal/mol for the cluster model and 28.4 kcal/mol for the QM/MM calculations. In INT3.2, the amine of the substrate is stabilized by the phenol proton of Tyr60 and the charges of the active center are again less polarized. For these reasons, both models have the same relative energy. As seen in Fig. 3, the cluster model yields a relative potential energy of -16.0 kcal/mol (we could not distinguish between INT3.1 and INT3.2 in the cluster model) and the QM/MM model has relative potential energies of -16.4 kcal/mol (INT3.1) and -18.4 kcal/mol (INT3.2).

The next step is the transfer of the Tyr60 proton to the nitrogen atom of the amine, now bonded to the cofactor. As a result of this proton exchange the amine group is again positively charged and the tyrosine becomes negatively charged. This leads again to energy differences between cluster and QM/MM models in TS4, and INT4. We can detect a specific interaction in INT4 that stabilizes the QM/MM model, between the amine group and the previously mentioned MM layer asparagine, through a H-bond chain made by three water molecules. This chain is stabilizing both the positive amine and the negative Tyr60. In order to include this effect in the cluster model one would have to model additional residues which would make the model too large and the calculations too slow.

These conclusions are maintained for TS5 and the products. The charges present in the active center make the cluster model more unstable in relation to the QM/MM model. Curiously, the activation

energy of TS5 for the cluster model is much smaller, 19.7 kcal/mol, vs 31.8 kcal/mol in the QM/MM model. We think this difference is justified by steric hindrance caused by the presence of the deaminated substrate in the catalytic pocket. Probably, at this stage, the product has already left the active center of the enzyme, which facilitates the exit of the amine group. If this is true, the energy of the QM/MM model is penalized in relation to the same structure of the cluster model, which has more space to accommodate the free amine group.

For the Single-Step mechanism, the geometries for the first minimum (R) and the transition state (TS) are very similar as well as the activation energy, with a difference of only 1.5 kcal/mol. At the end of the mechanism (P), the QM/MM model has the NH₃ interacting with the carboxylic oxygen atoms of substrate, whilst the cluster model has a less favorable conformation, which can explain the energy difference between both models.

In conclusion, the geometrical parameters obtained for the TAL reaction path with both models are very similar, especially those of the transition state structures. This result validates both methodologies, and makes us confident that the proposed reactions paths are well described. Certain conformations are more stable in the QM/MM model because they are stabilized by interactions with the rest of the enzyme, which is not represented in the cluster model. Our results show, as has been accepted for some time, that the cluster model approach is able to describe correctly the reaction path of enzymatic reactions, but that the obtained activation energies are usually higher due to the lack of stabilization of certain states by the enzyme. For the remaining sections, we will focus the discussion on the results of the QM/MM model.

Comparison between the two reaction paths

In order to choose the most probable reaction mechanism for tyrosine ammonia lyase one must ponder on three topics: the likelihood of the reactants state, the size of the barriers, and the consistency with mutagenic studies.

In relation to the likelihood of the reactants state, in order to calculate the exact energy of converting the N-MIO mechanism reactants into the Single-Step mechanism reactants, we would need to do some kind of free energy calculation like thermodynamic integration. In the absence of that information, and even without an energy value, it is easy to see that the reactants states of the Single-Step mechanism should be much more probable. The side chain of a Tyrosine has a pKa of 10: it is much more probable to find a protonated tyrosine amino acid than a deprotonated one in standard conditions. This predisposition should only be reversed if in the enzymatic vicinities of the tyrosine there were positive or partially positive amino acids (with a strong dipole, pointing in the opposite direction, for example). In the initial state of the N-MIO mechanism, it is required the active center of TAL to possess two deprotonated tyrosine residues. This is a very improbable event, not only because of the multiplicative relation between the two low probabilities, but also because the presence of a negative charge in the active center makes the presence of another negative charge yet more unlikely. Even if one considers the intermediate 1 structure, with only one deprotonated tyrosine, to be the reactants state (it is possible that the substrate enters the enzyme in the deprotonated amine form), it is very difficult to justify the deprotonation of Tyr60, since there are no residues in the vicinity to stabilize the negative charge. On the other side, the initial state of the Single-Step mechanism seems much more natural. For once, the overall charge of the active center is zero, whereas in the MIO mechanism it was -1 . Also, the Single-Step mechanism only requires one deprotonated tyrosine amino acid, Tyr300, and the negative side chain of this residue is stabilized by both the amine

group of Gln436, and by the electrophilic methylene group of the cofactor.

The obtained activation energies also favor the Single-Step mechanism. In the N-MIO mechanism, the limiting step in terms of kinetics is the third transition state (not considering the QM/MM energy of TS5 which could be attributed to steric hindrance which may be a limitation of the model). In relation to the initial reactants, TS3 has a very small barrier, 6.8 kcal/mol, but to calculate the rate limiting step one must use the energy of the previous intermediate, INT2, which is much more stable than the initial reactants. The real barrier that the reaction path must cross is 28.4 kcal/mol, it is the difference between the energy of TS3 and the energy of INT2. The Single-Step mechanism, has only one activation energy: 16.6 kcal/mol. This barrier is lower than the previous one and compatible with the experimental k_{cat} value.

Mutagenic studies showed that there are three residues essential to catalysis in TAL: Tyr60, Tyr300, and the MIO cofactor [12]. The three residues participate in both mechanisms, but in different roles. In the N-MIO mechanism the role of the residues is straightforward: all of them participate directly in the reaction path with breaking and formation of covalent bonds. Tyr300 accepts a proton from the amine group of the substrate in TS1, and then donates the same proton to the now free ammonia to make a positively charged ammonium ion. The MIO cofactor forms a covalent intermediate with the substrate during TS2, and then it frees an ammonia molecule in TS5. Finally, Tyr60 accepts a proton from the substrate in TS3, and donates it back to the amine group in TS4.

The roles of these three residues are not so strong in the Single-Step mechanism. Tyr300 is the only one that participates in the reaction directly by removing a proton from the substrate at the transition state. A putative role for Tyr60 is the stabilization of the leaving amine group, since the side chain phenol oxygen forms two hydrogen bonds with the leaving group. It may also be involved in stabilizing the orientation of the substrate. The cofactor does not participate directly in the reaction, also. It is possible that its electrophilic character helps the stabilization of the negative Tyr300 in the reactants state, making it more probable for the enzyme to be found in this state.

In conclusion, the initial state configuration and the activation energies favor the Single-Step mechanism, whereas the N-MIO mechanism is more consistent with the mutagenic experimental results. Additional studies are necessary to assert without doubts the reaction path followed by tyrosine ammonia lyase.

Comparison with the experimental k_{cat}

The turnover number for the reaction catalyzed by TAL is 27.7 s⁻¹ at 35° [1]. By using the transition state theory equation below, it can be calculated that this value corresponds to a barrier of 16.0 kcal/mol.

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

This k_{cat} is entirely compatible with the activation energy we obtained for the Single-Step mechanism (16.6 kcal/mol), but not with N-MIO mechanism (28.4 kcal/mol). Furthermore, the barrier of the N-MIO mechanism is not compatible with enzymatic catalysis in general, since it corresponds to a turnover number of 6 per year.

Comparison with a previous theoretical study

As we were developing our work in tyrosine ammonia lyase, another study was published that overlapped with some of the results we obtained [16]. We were happy to verify that, despite

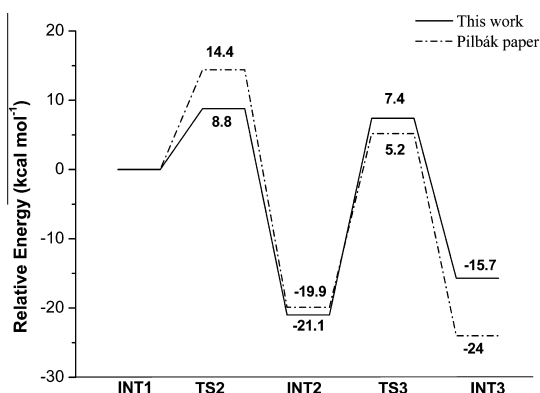


Fig. 6. Superposition of the energy profile presented in Pilbák work [17] (potential energy) for the N-MIO mechanism with the potential we obtained for the same stationary states with our QM/MM model (free energy). The nomenclature of the states is in respect with the present work.

the methodological differences, the results of both papers, where the subject was the same, are essentially identical. This result shows that both models are robust and reinforces the conclusions taken by both papers. Pilbák and colleagues used a 412 atoms ONIOM model with 114 atoms in the high layer at the B3LYP/6-31G(d):UFF level and mechanical embedding. In our model, we used ONIOM model with a high layer of 99 atoms and a total of 5216 atoms, at the B3LYP/6-311++G(2d,2p):AMBER// B3LYP/6-31G(d):AMBER level and electrostatic embedding. Among other things, Pilbák and colleagues studied five states of the N-MIO mechanism that correspond to our INT1, TS2, INT2, TS3, and INT3.1 states. The structures and relative energies obtained for those states are equivalent in both studies, as can be seen in Fig. 6.

In the initial geometry of Pilbák's work (INT1 in our nomenclature), the protonated tyrosine (Tyr300) establishes a hydrogen bridge with the amine (1.718 Å) in the same way as our QM/MM model (1.587 Å). The distance between the amine and the cofactor carbon atom, is also in the same range for both models (3.292 Å in Pilbák and 3.063 Å in our work). The first barrier, which corresponds to the nucleophilic attack of the amine group to the MIO cofactor has 8.8 kcal/mol in our study and 14.4 kcal/mol in the previous theoretical study. In this transition state, the amine has distanced itself from the Tyr300 and has come closer to the cofactor (2.873 Å in Pilbák work and 2.432 Å in ours). The H-bond between the amine and Tyr300 is still present but weaker and, in both models, the proton rotates in order to form a H-bond with one nitrogen in the imidazole ring of the cofactor. In both works, the proton of the substrate that will later bond with the Tyr60 is already interacting with it (2.232 Å in Pilbák work and 2.294 Å in ours). In both cases, this step is the most energetic, in relation to the reactants (INT1, in our nomenclature). The following intermediate (INT2) is very stable and with almost the same energy in both models (-21.1 kcal/mol in our model and -19.9 kcal/mol in Pilbák study). At this intermediate a covalent bond is formed between the substrate's amine and the cofactor. The proton from Tyr300 is now forming a H-bond with a nitrogen of the cofactor. In both studies, one proton of the amine is making a bridge with the oxygen atom of Tyr300, which stabilize the positive charge on the amine. The second transition state (TS3 in our nomenclature) is also equally energetic in both studies (7.4 in our model and 5.2 kcal/mol in Pilbák's work). This step shows a concerted breaking of the N-C (substrate) bond and deprotonation of the substrate by Tyr60. In the second transition state we see that the proton is already

completely bonded to the Tyr60 with a distance of 0.984 Å in our work, and 1.123 Å in the previous work. On the other hand the C-N bond is still breaking. In our work it elongated almost 0.55 Å, from 1.510 Å to 2.051 Å. In the previous work, the elongation was more subtle, changing from the intermediate distance of 1.532 Å to the transition state distance of 1.643 Å. The amine, at the end of the mechanism, is at a distance of 3.448 Å from the substrate in our work and at 2.597 Å in the previous one. There is also a difference in both models for the proton's distance to the substrate. In our work the Tyr60 suffers a rearrangement and the proton is now at 2.667 Å from the carbon while in Pilbák's work the proton lingers close to carbon, at 1.853 Å.

In conclusion, the results obtained in our work are very similar to those obtained by Pilbák and his colleagues, both concerning the geometries and the energetics of the mechanism. The activation energy of the TAL mechanism is given by the second compared step in both models, in our work it amounts to 28.5 kcal/mol, and to 26.3 kcal/mol in Pilbák's work. Although this problem is not addressed in Pilbák's paper, we think these values are a serious hindrance to the acceptance of the N-MIO mechanism. This potential energy surfaces mean that once the protein reaches the low energy intermediate, it is trapped for at least a few hours in that state. That result is not compatible with the TAL experimental kinetics or enzymatic catalysis in general.

Conclusions

The aim of this work was to understand the catalytic mechanism of the reaction catalyzed by tyrosine ammonia lyase. We built two models for this purpose: a QM model with 176 atoms, and a QM/MM model with 5216 atoms. With a cluster model we are able to describe more residues with QM, while maintain a tight control in the interactions that contributions to catalysis and in the movements of all residues. The QM/MM model with electrostatic embedding allows us to take into account the effect of the enzyme in the reaction at a QM level.

We explored two reaction paths, the N-MIO and the Single-Step mechanisms. Our results for the N-MIO mechanism are in agreement with previous studies and hypothesis concerning TAL. We studied two additional steps of this mechanism that were not previously described. The N-MIO mechanism appeal is that it clarifies very well the role of the three catalytic residues in the reaction: Tyr60, Tyr300, and the MIO cofactor. The three residues make covalent bonds or change protons with the substrate at some point in the reaction. Unfortunately, the activation energy calculated for this reaction path is rather high, 34.6|28.5 kcal/mol, and its reactants state, with two deprotonated tyrosine residues, seems implausible. The second reaction path we studied was not known or suggested previously. It consists in a single reaction step, where a proton from the substrate goes to Tyr300 and the amine group is released promptly. This step has an activation energy of 18.1|16.6 kcal/mol which is in agreement with the experimental value of 15.6 kcal/mol. The roles of Tyr60 and of the MIO cofactor are not so expressive in this mechanism, but they are still important for the reaction. The reaction is highly exothermic, which also agrees with the experimental observed irreversibility.

We think that further experimental and theoretical studies will be necessary to establish the catalytic mechanism of TAL and the related PAL and HAL enzymes. We hope the present work will be an important piece of the puzzle when that moment comes.

Acknowledgments

We would like to thank Professor László Poppe for providing us the structures of their model. A.J.M.R. thanks Fundação para a

Ciência e Tecnologia (FCT) for a doctoral scholarship (SFRH/BD/61549/2009). J.S.P. thanks to Dipartimento di Fisica di Unical for a doctoral scholarship.

References

- [1] J.A. Kyndt, T.E. Meyer, M.A. Cusanovich, J.J. Van Beeumen, *FEBS Lett.* 512 (2002) 240–244.
- [2] J.A. Kyndt, F. Vanrobaeys, J.C. Fitch, B.V. Devreese, T.E. Meyer, M.A. Cusanovich, J.J. Van Beeumen, *Biochemistry* 42 (2003) 965–970.
- [3] M. Berner, D. Krug, C. Bihlmaier, A. Vente, R. Müller, A. Bechthold, *J. Bacteriol.* 188 (2006) 2666–2673.
- [4] K.T. Watts, P.C. Lee, C. Schmidt-Dannert, *ChemBioChem* 5 (2004) 500–507.
- [5] Y. Zhang, S.-Z. Li, J. Li, X. Pan, R.E. Cahoon, J.G. Jaworski, X. Wang, J.M. Jez, F. Chen, O. Yu, *J. Am. Chem. Soc.* 128 (2006) 13030–13031.
- [6] Z. Xue, M. McCluskey, K. Cantera, A. Ben-Bassat, F.S. Sariaslani, L. Huang, *Enzyme Microb. Technol.* 42 (2007) 58–64.
- [7] T. Vannelli, Z. Xue, S. Breinig, W.W. Qi, F.S. Sariaslani, *Enzyme Microb. Technol.* 41 (2007) 413–422.
- [8] T.F. Schwede, J. Rétey, G.E. Schulz, *Biochemistry* 38 (1999) 5355–5361.
- [9] G.V. Louie, M.E. Bowman, M.C. Moffitt, T.J. Baiga, B.S. Moore, J.P. Noel, *Chem. Biol.* 13 (2006) 1327–1338.
- [10] M. Baedeker, G.E. Schulz, *Structure* 10 (2002) 61–67.
- [11] K.T. Watts, B.N. Mijts, P.C. Lee, A.J. Manning, C. Schmidt-Dannert, *Chem. Biol.* 13 (2006) 1317–1326.
- [12] A.C. Schroeder, S. Kumaran, L.M. Hicks, R.E. Cahoon, C. Halls, O. Yu, J.M. Jez, *Phytochemistry* 69 (2008) 1496–1506.
- [13] H. Ritter, G.E. Schulz, *Plant Cell* 16 (2004) 3426–3436.
- [14] M. Langer, A. Pauling, J. Rétey, *Angew. Chem. Int. Ed. Engl.* 34 (1995) 1464–1465.
- [15] L. Poppe, J. Rétey, *Angew. Chem. Int. Ed. Engl.* 44 (2005) 3668–3688.
- [16] J.D. Hermes, P.M. Weiss, W.W. Cleland, *Biochemistry* 24 (1985) 2959–2967.
- [17] S. Pilbák, Ö. Farkas, L. Poppe, *Chemistry* 18 (2012) 7793–7802.
- [18] L.L.C. Schrödinger, The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, 2010.
- [19] R. Dennington II, T. Keith, J. Millam, Semichem Inc., Shawnee Mission KS, 2009.
- [20] J.M. Wang, W. Wang, P.A. Kollman, D.A. Case, *J. Mol. Graphics Modell.* 25 (2006) 247–260.
- [21] D.A. Case, T.A. Darden, Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, M. Crowley, R.C. Walker, W. Zhang, K.M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, P.A. Kollman, AMBER 10, University of California, San Francisco, 2008.
- [22] J.M. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, *J. Comput. Chem.* 25 (2004) 1157–1174.
- [23] C.C.J. Roothaan, *Rev. Mod. Phys.* 23 (1951) 69–89.
- [24] C. Bayly, P. Cieplak, W. Cornell, P. Kollman, *J. Phys. Chem.* 97 (1993) 10269–10280.
- [25] C.T. Lee, W.T. Yang, R.G. Parr, *Phys. Rev. B* 37 (1988) 785–789.
- [26] A.D. Becke, *J. Chem. Phys.* 98 (1993) 5648.
- [27] R. Ditchfield, W.J. Hehre, J.A. Pople, *J. Chem. Phys.* 54 (1971) 724.
- [28] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, *J. Am. Chem. Soc.* 117 (1995) 5179–5197.
- [29] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian Inc., Wallingford CT, 2009.
- [30] T. Clark, J. Chandrasekhar, *J. Comput. Chem.* 4 (1983) 294–301.
- [31] M.J. Frisch, J.A. Pople, J.S. Binkley, *J. Chem. Phys.* 80 (1984) 3265.
- [32] A.D. Mclean, G.S. Chandler, *J. Chem. Phys.* 72 (1980) 5639–5648.

Establishing the Catalytic Mechanism of Human Pancreatic α -Amylase with QM/MM Methods

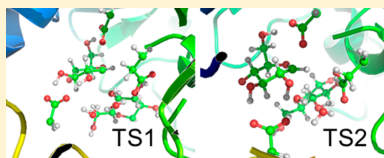
Gaspar P. Pinto,^{†,‡} Natércia F. Brás,[‡] Marta A.S. Perez,[‡] Pedro A. Fernandes,[‡] Nino Russo,[†] Maria J. Ramos,[‡] and Marirosa Toscano^{*,†}

[†]Università della Calabria, Via Pietro Bucci, Arcavacata, Italia

[‡]UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal

Supporting Information

ABSTRACT: In this work, we studied the catalytic mechanism of human pancreatic α -amylase (HPA). Our goal was to determine the catalytic mechanism of HPA with atomic detail using computational methods. We demonstrated that the HPA catalytic mechanism consists of two steps, the first of which (glycosylation step) involves breaking the glycosidic bond to culminate in the formation of a covalent intermediate. The second (deglycosylation step) consists of the addition of a water molecule to release the enzyme/substrate covalent intermediate, completing the hydrolysis of the sugar. The active site was very open to the solvent. Our mechanism basically differs from the previously proposed mechanism by having two water molecules instead of only one near the active site that participate in the mechanism. We also demonstrate the relevant role of the three catalytic amino acids, two aspartate residues and a glutamate (D197, E233, and D300), during catalysis. It was also shown that the rate limiting step was glycosylation, and its activation energy was in agreement with experimental values obtained for HPA. The experimental activation energy was 14.4 kcal mol⁻¹, and the activation energy obtained computationally was 15.1 kcal mol⁻¹.



INTRODUCTION

Human pancreatic α -amylase (HPA), an enzyme that belongs to the family of α -amylases [$\alpha(1-4)$ glucan-4-glucanohydrolase, EC3.2.1.1], promotes the hydrolysis of starch on $\alpha(1-4)$ glycosidic bonds. HPA belongs to the glycoside hydrolase (GH) family of enzymes, which catalyze the cleavage of

glycosidic bonds. This reaction, without the help of GH, would have a half-life of 5 million years. α -Amylases can be found in a large range of organisms. In humans, there are two types of α -amylase enzymes: one in the pancreas and the other in saliva. These two types differ in only 15 out of 496 amino acid residues, though some are in the active center.

The website <http://www.cazy.org>, which groups enzymes into families according to their sequence, has classified this α -amylase in the GH family number 13 within more than 100 families of enzymes described.

The α -amylases perform a wide variety of reactions involving glycosidic bonds from gly- and deglycosylation to hydrolysis and cyclization. Hyperactivity of HPA is connected to type II diabetes,¹⁻³ and controlling the enzyme is of paramount importance. Hence, knowing how the enzyme works is important for designing new inhibitors that are specific to HPA.

In GH enzymes, the mechanism is said to be either retaining or inverting, and the anomeric carbon (C₁) where the reaction takes place has two possible conformations, α and β . If the initial configuration is changed after glycosidic bond breakage, the mechanism is considered to be inverting; conversely, a retaining mechanism is that in which the anomeric carbon does not change its initial stereochemistry after hydrolysis is complete.

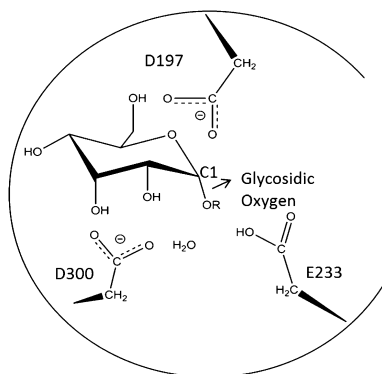


Figure 1. Quantum mechanics (QM) layer used in the enzyme/substrate conformation labeled for reference. Top to bottom: aspartate 197, anomeric carbon (C₁), glycosidic oxygen, aspartate 300, and glutamate 233. R = glucose ring.

Received: November 28, 2014

Published: May 13, 2015

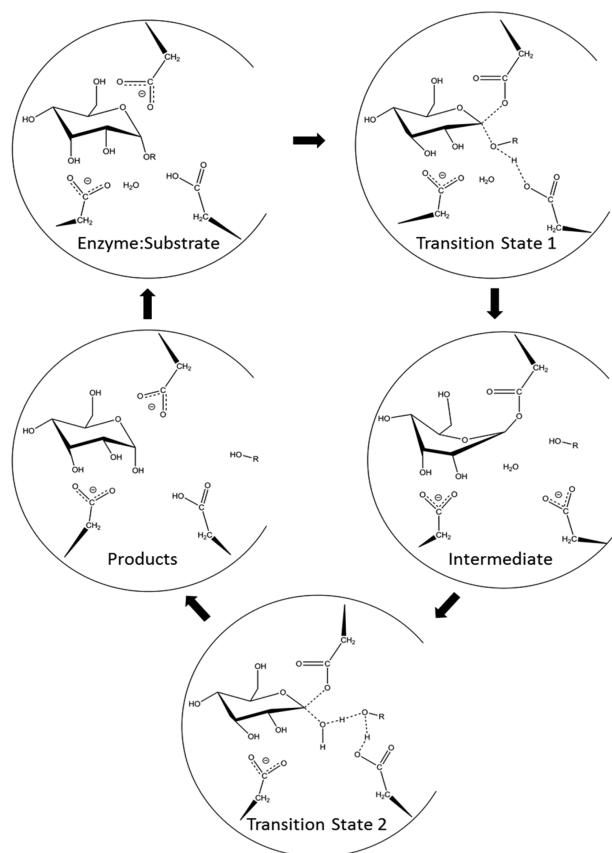


Figure 2. Proposed reaction mechanism for HPA. Residues and the substrate are in the same disposition as in Figure 1.

The general mechanism proposed for GH enzymes involves a couple of carboxylic acids (Asp and Glu residues). They catalyze the glycosylation step, thus yielding a covalent intermediate. This is followed by a nucleophilic attack from a water molecule on the anomeric carbon, forcing the covalent bond to break thus generating the deglycosylation step. Despite all of the mechanistic details already known from previous studies of other glycoside hydrolase enzymes,^{4–6} some points specific to the HPA mechanism have not yet been answered.

At the time we began our study, no structures of HPA with the natural substrate bound existed, so we determined an X-ray structure of human pancreatic amylase, the active site of which contained a docked inhibitor, which was a modified version of the commercial inhibitor Acarbose.

The amino acids that participate directly in the reaction mechanism were identified in the literature.⁴ Nevertheless, after further analysis of the active site and the interactions made by the protein with the modeled substrate, we identified one other aspartate that we found to be important in maintaining the substrate in the proper conformation for the mechanism to occur. In Figure 1, we present the active site and identify the most important amino acids.

Analyzing the tertiary structure of the enzyme, we found that the active site has a shallow pocket and that the solvent

availability was very high; thus, it was apparent that the number of water molecules that would participate in the reaction mechanism could be more than one to facilitate the second step.

The literature also tells us that one of the residues that participates actively in the mechanism should be protonated and that the other should be deprotonated, as shown in studies similar to our own, and in other mechanisms, such as those involving HIV-protease, where two amino acids in the same environment need different protonation states to initiate the mechanism.^{6–8}

Nevertheless, we decided to run a molecular dynamics (MD) simulation with both residues deprotonated. Because the active site is shallow, and with the solvent always near the substrate, it is possible that the mechanism could be initiated by a water molecule. We did not pursue this model any further with both residues deprotonated because residue E233 occupied the space near the glycosidic oxygen, blocking any water molecule that might enter the pocket deep enough to function as the proton donor (Figure 1).

Glycosidic linkages are enzymatically hydrolyzed by one of the two major groups of glycosidases: retaining and inverting.⁸ The retaining mechanism has been proposed for α -amylase,⁹ maintaining the initial conformation on the anomeric carbon.

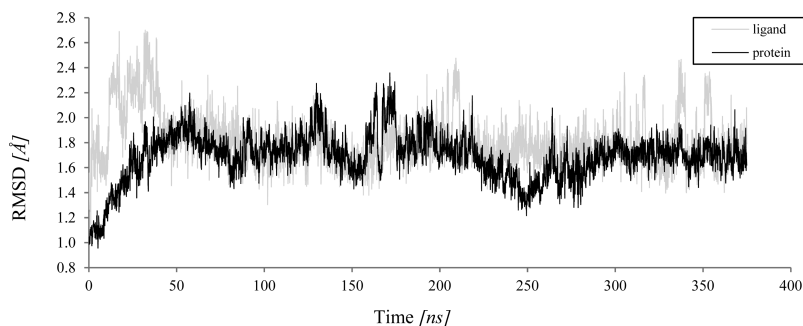


Figure 3. RMSD's of the enzyme/substrate complex, the substrate by itself, and the catalytic residues. Gray line = enzyme/substrate α carbons; black line = substrate.

In our case, however, instead of a sole pair of carboxylic acids around the anomeric carbon, there is also a third (another Asp) that stabilizes two water molecules present in the vicinity of the anomeric carbon (C_1). We have found that the presence of this second aspartate is very important; otherwise, the water molecule would be too far away to act as a nucleophile for the second mechanistic step. When starch interacts with HPA, a hydrogen bond is formed between Asp197 (nucleophilic residue) and the substrate, and the same happens with Glu233 (the proton donor group). For the transition state of the first step (TS1), it is believed that three events happen sequentially: the glycosidic bond breaks, the proton from Glu233 is transferred to the glycosidic oxygen, and anomeric carbon atom C_1 binds to nucleophile Asp197. In the end, at the intermediate geometry, there is a covalent interaction between an oxygen atom of Asp197 and the anomeric carbon. After the first mechanistic step, the water in the active site positions itself closer to C_1 , occupying the space that was previously occupied by the glycosidic oxygen. In the second mechanistic step, the water molecule promotes deglycosylation and simultaneously loses a proton that, through the glycosidic oxygen, reprotonates Glu233, as it had at the beginning of the mechanism (Figure 2).

METHODS

Models. The complex HPA/substrate was modeled from an HPA/acarbose X-ray crystallographic structure that was obtained from the PDB databank¹⁰ (ID: 1CPU⁵) with a resolution of 2.00 Å. The resulting model was used as a starting point for the computational studies. Molecular dynamics (MD) simulations were performed to evaluate the presence of water molecules near the active center and to obtain confirmation of the protonation state of the catalytic residues Asp197 and Glu233. Knowing that Asp197 was the nucleophile thus being deprotonated, two models were built with a protonated and deprotonated Glu233.

For these MD simulations, we used the amber force field (ff99sb)¹¹ and GLYCAM06¹² force field parameters for the proteins and the starch (substrate with 5 sugar rings bonded with 1–4 linkages), respectively. Explicit solvent TIP3P water in a cubic box with a minimum of 12 Å around the protein was used, and a sodium (Na^+) ion was added to neutralize the charge of the system. The resulting models consisted of 67932 and 67933 atoms for the protonated and deprotonated Glu233 systems, respectively.

Molecular Dynamics Simulations. Protonated and Deprotonated Models. For each model, we ran two

minimization steps: the first with the protein fixed and the water molecules free to move, and the second in which the whole system had no constraints. The minimized structures were then submitted to a 200 ps long warm-up simulation from 0 to 300 K at a constant volume and using periodic boundary conditions. Further, 370 ns Langevin dynamics with a 2 fs integration step were carried out. Nonbonded interactions were treated with the PME algorithm, and the cutoff was set to 10 Å, and the real part was also truncated at 10 Å with a pressure of 1 atm at 310 K with the NTP ensemble. All calculations were performed using the Amber12¹³ simulations package. Bonds involving hydrogen atoms were constrained with the SHAKE algorithm.¹⁴

Because the conformations obtained from the MD simulation for the deprotonated model showed the Glu233 far from the glycosidic oxygen and with no proton to initiate the mechanism, we ran QM/MM calculations with only the protonated model after a last minimization procedure.

QM/MM Calculations. QM/MM calculations were performed to determine the potential energy surface (PES) using Gaussian 09 software.¹⁵ The model used for the calculations was taken from the last minimization. The residues farther than 15 Å from the active site were frozen, and the water molecules from the solvent were cut off, yielding a model with a total of 7780 atoms and 505 residues. We divided our system into two layers, and the ONIOM formalism¹⁶ was used to calculate the corresponding PES. In the high layer, there were 75 atoms (i.e., two glycosidic rings, two water molecules, two aspartate side chains, and a glutamate side chain (until the β carbon)). The high layer was treated with density function theory (DFT) at the B3LYP/6-31G (d) level.^{17–21} The low layer (basically the rest of the enzyme together with the solvent water molecules) was treated at the molecular mechanics level with the parameters of the amber force field package. We used electrostatic embedding to treat the coupling between both layers.

To find transition states, we used flexible scans. The scans were made along the bond between the glycosidic oxygen and the acidic proton of Glu233 (first step) and along the bond between the water oxygen and the anomeric carbon of the covalent intermediate. Every residue within 15 Å of the active site was free, and the remaining residues were frozen. After the geometries of the transition states were roughly known, optimizations were carried out for the transition states. To ensure that the minima found corresponded to the same

reaction coordinates as the transition states, intrinsic reaction coordinate (IRC) calculations were made.

The zero point energies were calculated at the same QM/MM level as the geometry optimizations within the harmonic approximation. Thermal corrections (at 310 K) and entropies were calculated within the ideal gas/rigid rotor/harmonic oscillator model.

Single point energy calculations were performed subsequently with different density functionals (BB1K, M06, and B98),^{22–26} known to perform very well for thermodynamics and kinetics, together with the larger 6-311++G(2d,2p) basis set. Further dispersion calculations were carried for the B3LYP functional and M06. Charges were calculated through NBO analysis.^{27–32} H++ software was used to calculate the energetic cost of protonating the Glu233 in a biological environment.

RESULTS AND DISCUSSION

Because the initial file obtained from the PDB was an enzyme/acarbose complex, we had to model the natural substrate for

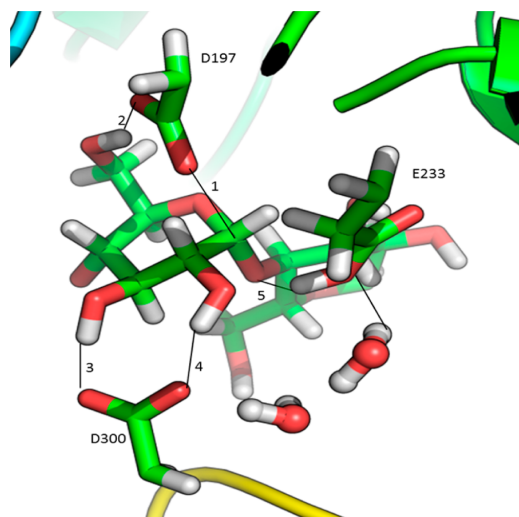


Figure 4. Reactant's three-dimensional structure of the active site with distances (Å) of the anomeric carbon to the nucleophilic residue.

HPA into the active site. As this was done manually, we performed an MD simulation to overcome any unfavorable interactions between the modeled substrate and the enzyme.

The structure of acarbose (active ingredient of some antidiabetic commercial drugs) has three glucose rings and a 4,5,6-trihydroxy-3-(hydroxymethyl) cyclohex-2-en-1-yl amine. The modified version has another glucose bonded to the cyclohexene. Being so similar to the actual substrate, the inhibitor that was already in place was logically used to model the substrate. Figure 1 shows that, even after the minimizations and molecular dynamics simulations, the modeled starch has a geometry very similar to that of the actual substrate. Certain differences are visible, as there should be because there are differences in the structures of both compounds, such as the presence of a nitrogen atom where the substrate has an oxygen atom, the absence of an oxygen atom in one cycle, and a hydroxyl group in another.

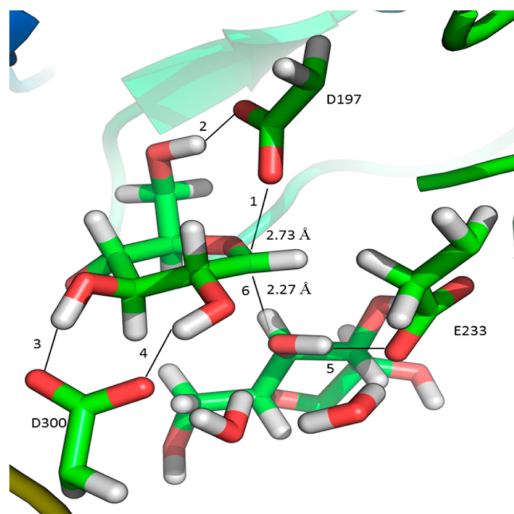


Figure 5. Transition state 1 three-dimensional structure of the active site with the distances (Å) of the anomeric carbon to the nucleophilic residue and the anomeric carbon to the glycosidic oxygen.

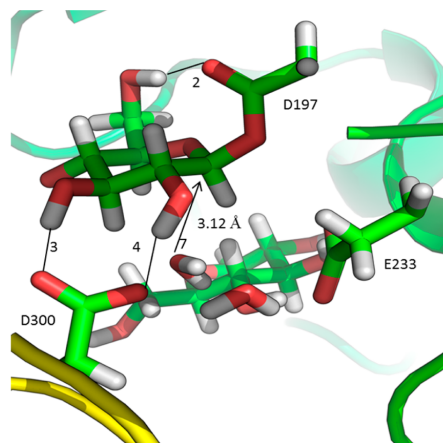


Figure 6. Intermediary with the oxygen (from water molecule)/anomeric carbon distance shown.

Obviously starch does not consist of only five glucose rings, it is a polymer that can have more five hundred monomers (glucose rings), but it is useless to simulate such a big substrate as only two of the glucose units are involved in the catalytic mechanism, so we modeled just 5 glucose rings.

After 350 ns of the production MD simulation, when the root-mean-square deviation (RMSD) for the whole enzyme was already stabilized and not higher than 1.5 Å at its highest point, we stopped the simulation.

As can be seen in Figure 3, the enzyme/substrate dynamics trajectory has RMSD values between 1.6 and 2.0 Å over time. The substrate has a higher deviation, but we confirmed that this was due to its sugar cycles oscillating conformations and the large mobility of the oxygen in the OH groups. One ring in

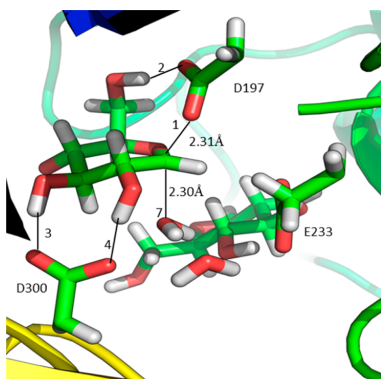


Figure 7. Second transition state with selected distances shown.

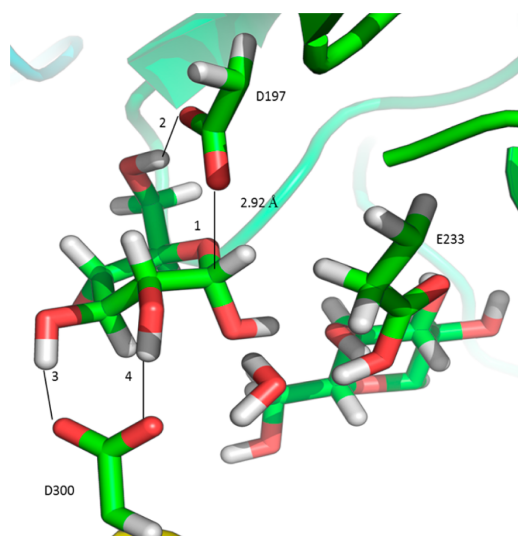


Figure 8. Products with selected distances shown.

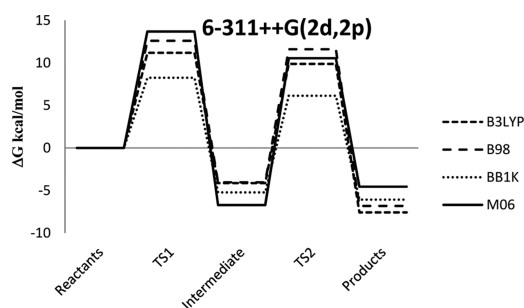


Figure 9. Potential energy surface for the HPA enzyme determined with 6-311++G(2d,2p) basis sets using four density functionals. The results were obtained from single point calculations.

particular, in position +3, has a conformational change at ~ 20 ns, but at 50 ns, it reverts back to its initial conformation. It is worth noting that this ring is in full contact with the solvent. We also noted that in the X-ray structure the temperature factors are higher in this area. The substrate interacts extensively with the side chains of amino acids in the active site, and therefore, it never leaves it during the simulation.

Analyzing the conformations obtained in the MD simulation, we saw that the substrate and its environment did not suffer big changes relatively to the initial conformation. Consequently, we proceeded to the QM/MM calculations with the conformation obtained in the last minimization procedure.

Glycosylation Step. After analyzing the reactant structure, we observed that the substrate in the active site was involved in four very strong hydrogen bonds; residue D300 established two hydrogen bonds (3 and 4 in Figure 4, represent 1.68 and 1.86 Å, respectively) with one of the sugar cycles, helping to keep it in the correct position for catalysis.

The E233 carboxylate also established two hydrogen bonds with the oxygen of the substrate, one as a donor atom (5 in Figure 4, represents 2.70 Å). The nucleophilic residue D197 looked like it was embracing the sugar ring that was going to lose the glycosidic oxygen (in Figure 4, it is making a hydrogen bond with residue E233).

One oxygen atom of D197 established a bridge with a hydroxyl group (2 in Figure 4, represents 1.81 Å), and the other oxygen atom was in a good position to attack the anomeric carbon (1 in Figure 4, represents 3.53 Å). The two water molecules that worked as a proton exchange network in the second step were positioned near the active site. Although they did not participate in the first step, they were near the active site and were stabilized between the D300 and E233 residues.

The HPA enzymatic mechanism is a retaining mechanism, meaning that the anomeric configuration of the substrate is maintained in the product. This means that the hydrolysis of the anomeric carbon should come from the alpha side of the anomeric carbon. The position of the water molecules, close to the residue Glu233, told us that those water molecules were important for the mechanism.

The proximity of the nucleophilic residue (D197) was of extreme importance for the first step to occur. It forced the anomeric carbon–glycosidic oxygen bond to break and the subsequent transfer of the proton from the E233 residue to the D197 glycosidic oxygen.

As the nucleophilic D197 residue attacked the anomeric carbon, the glycosidic bond began to extend until it reached a breaking point (6 in Figure 5) and carbon atom C1 started to bond with the aspartate (1 in Figure 5). Simultaneously, the proton bonded to the oxygen belonging to acid/base residue E233 (Figure 2) began to elongate. At this point, the anomeric carbon reached its breaking point. The proton immediately bonded to the glycosidic oxygen, as number 5 of Figure 5 shows, where the proton is already bonded to the glycosidic oxygen.

As a consequence of the D197 advance toward the anomeric carbon, the sugar ring had to adapt to the structural changes. While going from the reactant conformation to the TS1 conformation, the sugar ring changed from a stable chair conformation to a less stable half-chair conformation. Additionally, the anomeric carbon changed places until it reached a near planar angle with carbon C2 and the oxygen atom of the ring.

Table 1. Activation and Reaction Energies Obtained with Four Density Functionals and 6-311++G(2d,2p) as the Basis Set^a

6-311++G(2d,2p)	ΔG (ΔE) (kcal/mol)			
	TS1	intermediate	TS2	products
B3LYP	11.2 (14.0)	-4.1 (-3.2)	9.9 (12.1)	-7.6 (-5.9)
B98	12.6 (14.8)	-4.0 (-2.9)	11.6 (13.4)	-6.8 (-5.5)
BB1K	8.3 (9.8)	-5.2 (-4.0)	6.1 (7.4)	-6.1 (-5.3)
M06	13.7 (14.8)	-6.7 (-5.2)	10.5 (11.4)	-4.6 (-3.9)

^aThermal corrections and zero point energies were calculated for every stationary point and transition state. Electronic energies are shown in parentheses.

Table 2. Activation and Reaction Energies Obtained with Four Density Functionals and 6-31G(d) as the Basis Set^a

6-31G(d)	ΔG (kcal/mol)			
	TS1	intermediate	TS2	products
B3LYP	13.9	-8.8	8.0	-8.6
B98	15.2	-8.4	9.8	-7.9
BB1K	11.1	-8.8	4.8	-7.1
M06	17.7	-9.8	10.0	-5.1

^aThermal corrections and zero point energies were calculated for every stationary point and transition state.

NBO analysis gave us insight on point charges in every minima and transition state, thus confirming the dissociative nature of the transition state that was shown for similar enzymes in another study.⁴

The sugar ring changed back to the more stable chair conformation once the glycosylation intermediate was obtained. The anomeric carbon formed a covalent bond with the aspartate (see 1 in Figure 3 and Figure 6), yielding a very stable stationary state. The sugar ring was in its most stable conformation, the bond with D197 residue's oxygen was strong, and Figure 6 shows also the hydrogen bond established between the 6'-hydroxyl group of the sugar and one oxygen atom of the D197 residue (2 in Figure 6, represents 2.01 Å), which already existed at the beginning of the reaction. This H-bridge promoted the formation of a cycle with nine atoms with the weakest link being this hydrogen bridge between an OH group and the D197 residue.

On the alpha side of the sugar, one of the water molecules came close to where the leaving sugar was before (7 in Figure 4). This forced the leaving group to move farther away from its initial position. The second water molecule established H-bridges with residues E233 and D300.

Deglycosylation Step. In the second step of this reaction mechanism, the type of the reaction (retaining or inverting) is defined. HPA was already specifically known to be a retaining mechanism, meaning that the anomeric carbon maintains its initial isomeric configuration. This occurs when the water molecule attacks the anomeric carbon (7 in Figure 4) on the alpha side of the sugar ring.

In Figure 4, the water molecule and the D197 residue, resulting from the first mechanistic step, are on opposite sides

of the sugar ring. D197 is in a beta position, and the water molecule is in an alpha position.

These two residues, the water and the aspartate, with the anomeric carbon in the middle, form an angle of 140°. Chemically speaking, this kind of S_N2 reaction is more favorable with an angle of 180°, but in a biological medium such an angle becomes too difficult to achieve because of steric constraints around the reaction site.

Interestingly, the first mechanistic step is also an S_N2 reaction, but the angle (159°) formed by the glycosidic oxygen, the anomeric carbon, and the aspartate is closer to the ideal angle for these reactions.

From this stable stationary state, the mechanism continued with two events happening almost simultaneously: as the oxygen of the water molecule approached the anomeric carbon, both of its hydrogen-oxygen bonds extended. Concomitantly, the covalent bond between the anomeric carbon and the aspartate started to break, initiating deglycosylation. As Figure 7 shows, the transition state was reached when the anomeric carbon was halfway between the water and the aspartate. As with the first mechanistic step, the sugar ring twisted its preferred chair conformation into a half-chair. At that moment, one proton from the water molecule broke away from the oxygen atom (8 in Figure 7) and approached the second water molecule, and consequently one proton of the second water molecule moved toward the E233 residue.

After the second transition state, the sugar ring adopted a stable chair conformation with a hydroxyl group bonded to the anomeric carbon in an alpha position, thus retaining the initial conformation (Figure 8). The second water molecule reprotonates the glutamate in the active site.

Having achieved a computational mechanism similar to that proposed for HPA and similar to other glycosidase enzymes, we compared the experimental value for the energy barrier with the energy barrier obtained in this work. Previously described, the experimental activation energy of the HPA mechanism is 14.4 kcal mol⁻¹.³³

In Figure 9, we present the single point calculations with four hybrid GGA density functionals with a triple- ζ basis set with diffuse orbitals (values are shown in Table 1). Frequencies for every stationary point and transition state were computed; thus, the results presented have zero point and thermal corrections. We see in Figure 9 that the activation energies obtained for each functional are lower than the experimentally obtained

Table 3. Activation and Reaction Energies Obtained with and without the Insertion of DFT-D3 Interactions

	ΔG (kcal/mol)			
	TS1	intermediate	TS2	products
B3LYP-D3 (B3LYP)	10.0 (11.2)	-4.2 (-4.1)	8.9 (9.9)	-3.8 (-7.6)
M06D3 (M06)	13.0 (13.7)	-6.3 (-6.7)	10.6 (10.5)	-3.4 (-4.6)

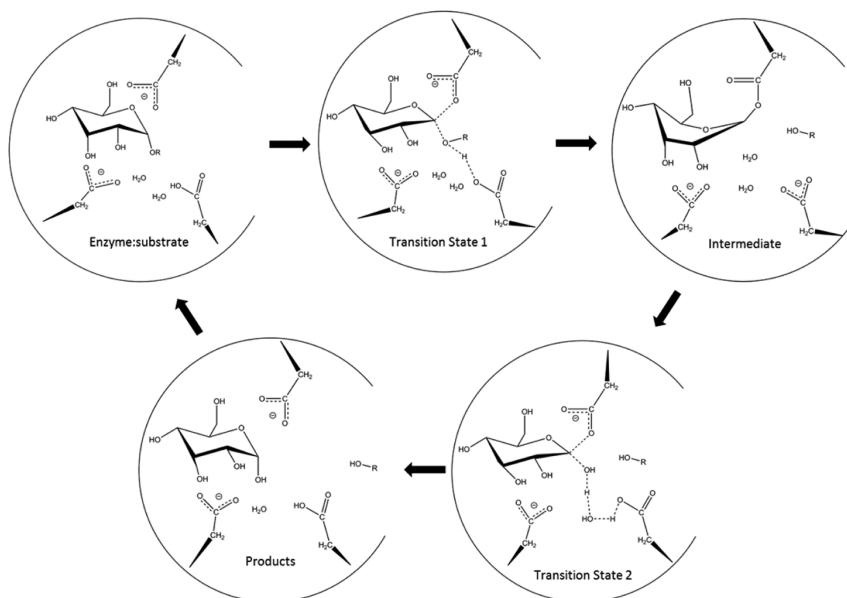


Figure 10. Mechanism obtained computationally in this work.

activation energy for this enzyme but are still within the error range given by the method. The same calculations were made with a lower level basis set as shown in Table 2. With a double- ζ basis set without diffuse orbitals, the activation energy rises for every density functional. Because the active site of our enzyme has various oxygen atoms in both carboxyl and hydroxyl groups, as well as two water molecules, a destabilization of the energy in the transition states was expected.

Analyzing the PES of both basis sets, we notice that even though TS1 has a higher ΔG , the rate limiting step is TS2 because of the intermediate stability. This is also in agreement with other glycosidase mechanisms. Both hybrid functionals yielded similar ΔG values for the activation energy with the lower basis set yielding slightly higher results.

Dispersion interactions were treated with DFT-D3 for B3LYP and M06, the results and how they compare with the results obtained without the insertion of DFT-D3 are shown in Table 3. The first transition state is stabilized, and the product energy is elevated. The intermediate has a very small elevation in the energy, and only in the second transition state is there a noticeable difference in the results obtained for both density functionals. In B3LYP, the second transition state is stabilized as it was for the first transition state; however, the second transition state ends up being slightly destabilized.

CONCLUSIONS

Despite the number of mechanistic studies performed on the glycosidase family of enzymes, a vast number of them are not understood or known in atomistic detail and require further study. In our work, we carried out a QM/MM (DFT:AMBER) study on the full enzyme to fully understand the catalytic mechanism and solvent availability near the active site. The active site has a high solvent availability, which supports our catalytic mechanism with two water molecules instead of only

one for the proposed mechanism. As is generally accepted for glycosidase enzymes, our analysis of the minima and transition states for this reaction mechanism yields a dissociative character for the transition state of the glycosylation step: the bond between the anomeric carbon and the glycosidic oxygen was at a breaking point at 2.26 Å; however, the glycosylation bond was still far (2.73 Å) from being formed.

In the second transition state, the water molecule was still not activated even though the anomeric carbon was midway between the previous glycosidic linkage and the water molecule's oxygen. Both protons had elongated bonds, but they were still both clearly bonded to the nucleophilic oxygen. In fact, only when one proton broke away was the oxygen from the former water molecule in a strong enough position to bond with the anomeric carbon, hydrolyze the sugar ring, and thus finish the second mechanistic step.

The reprotonation of the acid/base residue occurred through a water molecule. This is a change from the literature-proposed pathway for HPA (Figure 10). The activation energy is in total agreement with the experimental values in which the energy barrier is 14.4 kcal mol⁻¹.

DFT calculations yielded an activation energy of 11.2 kcal/mol. With the addition of the energetic cost of protonating Glu233, the activation energy rises to 15.1 kcal/mol, which is still in agreement with the experimental values.³³

ASSOCIATED CONTENT

Supporting Information

Superposition of the original inhibitor with the modeled substrate, superposition of the inhibitor with the modeled substrate in the conformation that we began in this study, a table with distances obtained through IRC of the first transition state, where we see clearly that TS1 is closer to the ES geometry than the intermediate geometry. The Supporting Information is

available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00222.

AUTHOR INFORMATION

Corresponding Author

*E-mail: marirosa.toscano@unical.it.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has been funded by FEDER/COMPETE and Fundação para a Ciência e a Tecnologia through Grants PEst-C/EQB/LA0006/2011 and EXCL/QEQ-COM/0394/2012. N.F.B. thankfully acknowledges her Fundação para a Ciência e Tecnologia grant (IF/01355/2014)

REFERENCES

- (1) deSales, P. M.; deSouza, P. M.; Simeoni, L. A.; Magalhaes, P. D. O.; Silveira, D. α -Amylase inhibitors: a review of raw material and isolated compounds from plant source. *J. Pharm. Pharm. Sci.* **2012**, *15*, 141–183.
- (2) Funke, I.; Melzing, M. F. Traditionally used plants in diabetes therapy – phytotherapeutics as inhibitors of α -amylase activity. *Rev. Bras. Farmacogn.* **2006**, *16*, 1–5.
- (3) Jayaraj, S.; Suresh, S.; Kadeppagari, R. K. Amylase inhibitors and their biomedical applications. *Starch/Staerke* **2013**, *65*, 535–542.
- (4) Bras, N. F.; Fernandes, P. A.; Ramos, M. J. QM/MM Studies on the α -Galactosidase Catalytic Mechanism: Hydrolysis and Transglycosylation Reactions. *J. Chem. Theory Comput.* **2010**, *6*, 421.
- (5) Passos, O.; Fernandes, P. A.; Ramos, M. J. QM/MM Study of the Catalytic Mechanism of GalNAc Removal from GM2 Ganglioside Catalyzed by Human β -Hexosaminidase A. *J. Phys. Chem. B* **2011**, *115*, 14751–14759.
- (6) Bras, N. F.; Ramos, M. J.; Fernandes, P. A. DFT studies on the β -glycosidase catalytic mechanism: The deglycosylation step. *J. Mol. Struct.: THEOCHEM.* **2010**, *946*, 125–133.
- (7) Brik, A.; Wong, C. H. HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.* **2003**, *1*, 5–14.
- (8) Chiba, S. A Historical Perspective for the Catalytic Reaction Mechanism of Glycosidase; So as to Bring about Breakthrough in Confusing Situation. *Biosci. Biotechnol. Biochem.* **2012**, *76*, 215–231.
- (9) Zhang, R.; Li, C.; Williams, L. K.; Rempel, B. P.; Brayer, G. D.; Withers, S. G. Directed “in Situ” Inhibitor Elongation as a Strategy To Structurally Characterize the Covalent Glycosyl-Enzyme Intermediate of Human Pancreatic α -Amylase. *Biochemistry* **2009**, *48*, 10752–10764.
- (10) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. E., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535–542.
- (11) Brayer, G. D.; Sidhu, G.; Maurus, R.; Rydberg, E. H.; Braun, C.; Wang, Y.; Nguyen, N. H.; Overall, C. M.; Withers, S. G. Subsite Mapping of the Human Pancreatic α -Amylase Active Site through Structural, Kinetic, and Mutagenesis Techniques. *Biochemistry* **2000**, *39*, 4778–4791.
- (12) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2000**, *65*, 712–725.
- (13) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Outeiro, J. G.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **2008**, *29*, 622–655.
- (14) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolosváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A. Kollman, P. A. *AMBER 12*; University of California: San Francisco, CA, 2012.
- (15) Hammonds, K. D.; Ryckaert, J. P. On the convergence of the SHAKE algorithm. *Comput. Phys. Commun.* **1991**, *62*, 336–351.
- (16) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, M. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., *Gaussian 09*, revision D.01; Gaussian Inc.: Wallingford, CT, 2009.
- (17) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A New ONIOM Implementation in Gaussian 98. Part I. The Calculation of Energies, Gradients and Vibrational Frequencies and Electric Field Derivatives. *J. Mol. Struct.: THEOCHEM* **1999**, *462*, 1–21.
- (18) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (19) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (20) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (21) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (22) Becke, A. D. Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, *104*, 1040–1046.
- (23) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (24) Schmider, H. L.; Becke, A. D. Optimized density functionals from the extended G2 test set. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
- (25) Foster, J. P.; Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
- (26) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Development and Assessment of a New Hybrid Density Functional Method for Thermochemical Kinetics. *J. Phys. Chem. A* **2004**, *108*, 2715–2719.
- (27) Reed, A. E.; Weinhold, F. Natural bond orbital analysis of near-Hartree–Fock water dimer. *J. Chem. Phys.* **1983**, *78*, 4066–4073.
- (28) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural-population analysis. *J. Chem. Phys.* **1985**, *83*, 735–746.
- (29) Reed, A. E.; Weinhold, F. Natural Localized Molecular Orbitals. *J. Chem. Phys.* **1985**, *83*, 1736–1740.
- (30) Carpenter, J. E.; Weinhold, F. Analysis of the geometry of the hydroxymethyl radical by the different hybrids for different spins natural bond orbital procedure. *J. Mol. Struct.: THEOCHEM* **1988**, *46*, 41–62.
- (31) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899–926.

(32) Weinhold, F.; Carpenter, J. E. *Electronic Structure-Theory. The Structure of Small Molecules and Ions*, 1st Edition; New York, NY, 1988; Vol. 1, pp 227–236.

(33) Rydberg, E. H.; Li, C.; Maurus, R.; Overall, C. M.; Brayer, G. D.; Withers, S. G. Mechanistic Analyses of Catalysis in Human Pancreatic R-Amylase: Detailed Kinetic and Structural Studies of Mutants of Three Conserved Carboxylic Acids. *Biochemistry* **2002**, *41*, 4492–4502.

Enzyme Catalysis

Triesterase and Promiscuous Diesterase Activities of a Di-Co^{II}-Containing Organophosphate Degrading Enzyme Reaction MechanismsMarta E. Alberto, Gaspar Pinto, Nino Russo, and Marirosa Toscano^{*,[a]}

Abstract: The reaction mechanism for the hydrolysis of trimethyl phosphate and of the obtained phosphodiester by the di-Co^{II} derivative of organophosphate degrading enzyme from *Agrobacterium radiobacter* P230 (OpdA), have been investigated at density functional level of theory in the framework of the cluster model approach. Both mechanisms proceed by a multistep sequence and each catalytic cycle begins with the nucleophilic attack by a metal-bound hydroxide on the phosphorus atom of the substrate, leading to the cleavage of the phosphate-ester bond. Four exchange-correlation functionals were used to derive the potential energy profiles in protein environments. Although

the enzyme is confirmed to work better as triesterase, as revealed by the barrier heights in the rate-limiting steps of the catalytic processes, its promiscuous ability to hydrolyze also the product of the reaction has been confirmed. The important role played by water molecules and some residues in the outer coordination sphere has been elucidated, while the binuclear Co^{II} center accomplishes both structural and catalytic functions. To correctly describe the electronic configuration of the d shell of the metal ions, high- and low-spin arrangement jointly with the occurrence of antiferromagnetic coupling, have been herein considered.

Introduction

Organophosphate triesters (OPs) do not occur naturally and have appeared in the environment in recent times with the widespread use of pesticides dating back to about 70 years ago. The toxicity of organophosphate triesters lies in the irreversible inhibition of acetylcholinesterase (AChE), preventing nerve signal transduction and causing the death of affected organisms.^[1] Their powerful anticholinesterase activities led to the further development of a range of nerve-gas agents and the consequent drive for efficient antidotes to organophosphorus poisons.

To deal with this problematic situation created by human interference, some bacteria have used environmental OPs contamination to their advantage by evolving enzymes able to degrade many of these compounds as a source of phosphorus for their metabolism.^[2–4]

The bacterial phosphotriesterases (EC 3.1.8.1), able to promote such hydrolysis processes, are members of the binuclear metallohydrolases, a large family of enzymes which require two metal ions in their active site to promote a plethora of hydrolytic reactions.^[5]

Because of the potential use of phosphotriesterases in the detoxification of organophosphate pesticides, there is considerable interest in better understanding of their catalytic mechanism and improving their efficiency.

A promiscuous enzyme (OpdA) capable of degrading a wide range of organophosphates and nerve agents in less or non-toxic compounds, has been isolated from an *Agrobacterium radiobacter* P230 strain (OPDA).^[6] This enzyme shows a high sequence identity (>90%) with closely related phosphotriesterases OPH first isolated from *Pseudomonas diminuta*,^[7] with *Flavobacterium* sp. ATCC 27551^[8] and with the highly promiscuous phosphodiesterase GdpQ from *Enterobacter aerogenes*.^[9]

While the mechanism of phosphotriester hydrolysis by OPH has been extensively studied,^[10–12] less is known about the same reaction catalyzed by OpdA. They share an identical set of metal-coordinating ligands but the reaction mechanism they employ may be different. The mechanisms proposed for the hydrolysis of the phosphate-ester bond catalyzed by this class of enzymes often involve a nucleophilic attack of a metal-bound hydroxide on the phosphorus atom of the substrate, leading to the cleavage of the phosphate-ester bond and the release of –OR group.^[5,13] Usually, the nucleophilic agent is coordinated to one or both metal centers in the active site. However, there are rare cases where the nucleophile is a non-coordinated hydroxide in the outer metal coordination sphere.^[5]

Like GdpQ, OpdA has been suggested to be promiscuous not only with respect to the substrate but also with respect to the metal ion composition.^[13] Atomic absorption spectroscopy and anomalous scattering measurements^[14,15] indicate that the

[a] Dr. M. E. Alberto, G. Pinto, Prof. N. Russo, Dr. M. Toscano
Dipartimento di Chimica e Tecnologie Chimiche and Centro di Calcolo ad Alte Prestazioni per Elaborazioni Parallele e Distribuite-Centro d'Eccellenza MIUR
Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)
E-mail: m.toscano@unical.it

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201405593>.

OpdA enzyme may contain a bimetallic centre of Fe^{2+} – Zn^{2+} in its native form, while some *in vitro* studies have shown that binuclear centres of Co^{2+} , Mn^{2+} and Cd^{2+} have a higher catalytic power than the native form of the enzyme.^[14,16,17]

The three-dimensional structure of OpdA is that of a homodimer the α and β subunit folds of which form a TIM barrel. The binuclear center of the enzyme is located in the C terminus of the β barrel. In the absence of substrate, the metal ions, referred to as α and β , are bridged by a carboxylated lysine residue (Lys169) and one bridging water/hydroxide. The α -metal ion is further coordinated to the side chains of His55, His57 and Asp301, while the β -metal ion is coordinated to the side chains of His201, His230 and two hydroxide/water ligands, appearing accordingly, more solvent-exposed.^[18–20]

Only three amino acid variations are observed in the substrate binding pocket between OpdA and OPH, namely Arg254His and Tyr257His, which are extensively involved in a network of hydrogen bonds and which play a role in modulating the OpdA-catalyzed reaction. Disruption of such networks has been shown to lead to OpdA forms with properties similar with OPH.^[13]

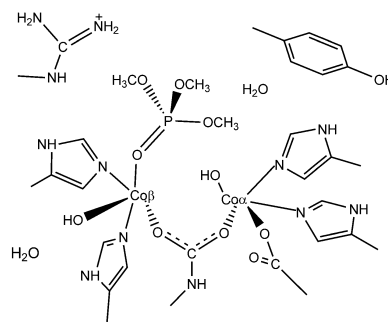
OpdA derivatives with di- Zn^{2+} and di- Cd^{2+} metallic center have been suggested to employ a mechanism similar to that of OPH having the same metal ion composition. For them only one kinetically relevant protonation equilibrium was observed in ethyl paraoxon hydrolysis at low pK_a range around 4–5, which is typical for a μ -OH nucleophile. On the contrary, the di- Co^{2+} derivative, the most studied OpdA form, may behave differently. Actually, depending on the metal ion substitution in OpdA, the latter displays a high mechanistic flexibility for OP degradation.^[17] For the Co^{II} derivative of OpdA two protonation equilibria ($\text{pK}_{a1} \sim 5$; $\text{pK}_{a2} \sim 10$) were identified as being relevant for catalysis, allowing the enzyme to reach the maximum catalytic efficiency at pH above 10. In this case a terminal hydroxide was suggested to act as nucleophile. Actually, analysis of the crystal structures of the Co-containing OpdA complex with inhibitor, substrate analogues^[18,21] or substrates were able to provide good models for the Michaelis complex,^[22] and together with the results of a molecular dynamics docking study with the substrate EPO (diethyl 4-methoxyphenyl phosphate),^[22] suggested a mechanism in which the nucleophile seems to be a hydroxide group bound to the α -metal ion. Such a nucleophile has been suggested to coordinate the α - Co^{2+} to saturate a vacant coordination site provided by a shift of the μ -OH toward the β - Co^{2+} , hypothesized in view of the observed reduction of the exchange coupling between metals upon substrate binding.^[23] The proposed shift should replace one of the two water molecules originally present at the β - Co^{2+} ion while the coordination of the substrate at the same site has been suggested to replace the other water molecule. This scheme is in agreement with the hypothesis of Ely et al.^[23] who stated that the number of ligands of β - Co^{2+} center reduces from 6 to 5 upon substrate binding, leaving unchanged the coordination number of the α -metal ion.

In the present study we have chosen to explore the action mechanism of a binuclear Co^{2+} OpdA enzyme in hydrolyzing a phosphotri- and the obtained phosphodiester, in order to

evaluate the promiscuous nature of this enzyme. On the basis of some preliminary computations, a QM cluster model has been constructed containing the di- Co^{II} metal center, their first coordination shell (six amino acid residues and two terminal coordinated OH), an arginine (Arg254) and a tyrosine (Tyr257) and two other water molecules lying in the outer metals coordination sphere.

Computational details

The cluster model used to simulate the active site of OPDA was constructed using conventional modelling of the amino acid side chains starting from the available X-ray structure [PDB code 2D2H]^[18] (Scheme 1).



Scheme 1. Model of the active center of the OPDA enzyme with trimethyl phosphate as substrate.

The first coordination sphere of the two Co^{2+} cations involves two $-\text{OH}$ groups, an anionic CH_3COO^- group to simulate the Asp301, four imidazole rings to mimic the amino acids His55, His57, His201, His230, and a $\text{CH}_3\text{NHCOO}^-$ charged group to model the carboxylated Lys169. Two water molecules, a $\text{CH}_3\text{NHC}(\text{NH}_2)_2^+$ group as simplified model of Arg254 and a protonated phenol to simulate a tyrosine residue Tyr257, were added in the second coordination sphere of enzyme. Since the PDB structure corresponds to the enzyme with the potential substrate bound non-productively,^[18] drawing upon this ion, we modelled a trimethyl phosphate coordinated to the β - Co^{II} . The final model has 122 atoms and a total charge of +1 (Figure 1). It is noteworthy that a series of preliminary calculations were performed to shed light on some crucial points concerning the cluster model, such as the protonation state of the hydroxide bound to the $\text{Co}-\beta$ metal ion and the preferred substrate coordination to $\text{Co}-\beta$ or $\text{Co}-\alpha$ (see the Results and Discussion section).

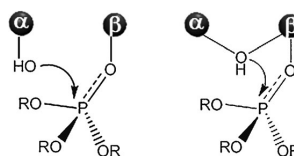


Figure 1. Initial proposals for the mechanism of OpdA enzyme.

All the calculations were performed using density functional theory (DFT), choosing for the optimizations in gas phase, the B3LYP^[24] functional which includes the Becke's hybrid exchange and correlation functional of Lee, Yang and Parr as implemented in the Gaussian 09 package.^[25] LANL2DZ^[26] pseudopotential and 6-31G(d,p)^[27] basis sets were used to describe transition metal ions and all the rest of elements, respectively.

Geometry optimizations, were followed by vibrational analysis on all stationary points, to evaluate their character as minima or saddle points.

Despite the performances of the B3LYP functional in describing the mechanisms involved in enzymatic catalysis is well supported by the literature,^[28,29] it is known that new developed functionals represent a better tool for the calculation of barrier heights, thermochemical kinetics and non-bonded interactions.^[30,31] As a consequence, in order to obtain more accurate energies in the protein environment, single-point calculations were performed on the optimized geometries at the B3LYP,^[24] B3LYP-D^[32] MPWB1K,^[33] and M06L^[33] levels using the larger basis set 6-311+G(2d,2p) in the framework of the conductor-like self-consistent reaction field polarized continuum model (CPCM).^[34,35] A dielectric constant of $\epsilon = 4$, accounting for the average effect of both the protein and the water medium surrounding the protein, was chosen to describe the protein environment.^[28]

Previous studies, also including charged systems, demonstrated that geometry relaxation effects obtained in solvent are usually rather small, as a consequence only single-point calculations on optimized structures were carried out to correct the corresponding free energy for the gas phase by the solvation energy, being quite sure not to commit significant errors.^[30c]

The most stable electronic configuration of the cluster was evaluated performing a series of single-point energy calculations on the enzyme-substrate model with different spin multiplicities. In particular values of $(2S+1)$ equal to singlet, triplet, quintet and septet were tested.

The occurrence of antiferromagnetic (AFM) coupling was also evaluated by defining a SCF guess with a multiplicity of +4 for one Co^{+2} and a multiplicity of -4 for the other metal ion, to simulate the high-spin configuration of each metal. The AFM coupling between the metal ions with low-spin configuration was also considered defining a SCF guess with a multiplicity of +2 for one Co^{+2} and a multiplicity of -2 for the other one.

Anyway, the trend established by the obtained energy values indicates that both cobalt ions favor the high-spin electronic configuration, resulting, the septet state, the most stable one.

Contrary to what was previously observed in other phosphatase enzymes,^[31a] the AFM coupling between the metal ions does not have any influence on the energetic profile of the reaction, this evidence being consistent with the weak ferromagnetically coupled Co^{II} ions observed experimentally on OpdA derivative (Table S1 in the Supporting Information).

Results and Discussion

The mechanism employed by OpdA to hydrolyze a range of OP substrates has not been investigated to the same extent as OPH, but metal-ion replacement and magnetic circular dichroism (MCD) studies^[23] have led to the proposal of two alternative mechanisms which can be modulated by both pH and metal-ion composition. In the first one, the metal ion bridging hydroxide is the proposed nucleophile to initiate hydrolysis, in

analogy to what has been previously found for purple acid phosphatase and other phosphatases.^[31,36] On the other one, a hydroxide terminally coordinated to the α -metal ion has been proposed as nucleophile^[36] (Figure 1).

As already described above, a combination of physicochemical studies on di- Co^{II} derivative^[17,23] suggests that the enzyme works better at higher pH ($\text{pH} > 10$) in which two relevant protonation equilibria have been detected, and the mechanism has been suggested to initiate with a nucleophilic attack by a terminally coordinated hydroxide. Although the high pH value at which experiment finds the maximum catalytic efficiency suggests a terminal hydroxide bound to β - Co^{II} , we have also considered the possibility to have a water molecule in place of a hydroxide. Our computations reveal that a very high activation barrier ($\approx 43.0 \text{ kcal mol}^{-1}$) is required in such conditions for the reaction to take place, demonstrating that the process proceeds at higher pH in which the ligand is an -OH group (Figure S1 in the Supporting Information).

Moreover, in order to evaluate the preferential coordination site for the trimethyl phosphate substrate, we have performed calculations exploring the two possible cases: coordination on α - Co^{II} and coordination on β - Co^{II} .

Computations have shown, beyond a shadow of a doubt, that coordination on the α center does not lead to a stable adduct, as the substrate, during optimization, moves away from the metal ion by up to 5.74 Å. This is in agreement with the experimental suggestion that indicates the coordination on α - Co^{II} as a non-productive beginning for the catalysis.

So, all the following results will be referred to as the model having a hydroxide as Co - β ligand and the substrate coordinated to the same β - Co^{II} .

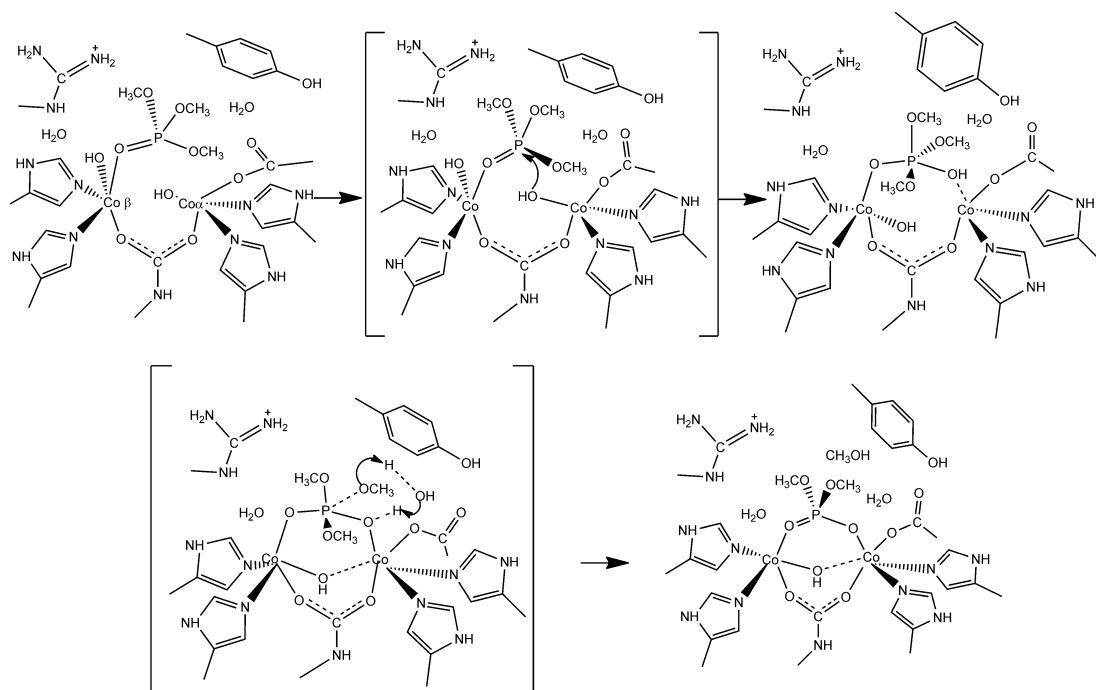
The proposed reaction mechanisms for the hydrolysis of both trimethyl- and dimethyl phosphate by di- Co^{II} -containing OpdA enzyme are depicted in Scheme 2 and Scheme 3.

The optimized structures of all the stationary points with structural key parameters can be found in Figure 2 and Figure 4. The corresponding potential energy profiles obtained in protein environment, for the most stable spin multiplicity $(2S+1=7)$, are shown in Figure 3 and Figure 5. Three XC-functionals have been used to derive the reported energetic values.

Trimethyl phosphate hydrolysis

The comparison between the optimized ES with the X-ray structure shows a quite good agreement of geometrical parameters (Table S2 in the Supporting Information).

The overall geometry of the metal centers as predicted by our computation is of a distorted square pyramidal geometry for the α - Co^{II} cation, which binds the lateral chains of His55, His57, Asp301, Lys169 and the plausible nucleophile OH, while the β - Co^{II} cation results in a quasi-octahedral geometry being coordinated to an oxygen atom of the substrate, His201, His230, Lys169, a hydroxide molecule and interacting with the nucleophile terminally coordinated to the α -metal. Actually, during the optimization, the latter orients itself to interact with both metals in a pseudo-bridging position (1.961 and 2.060 Å



Scheme 2. Proposed reaction mechanism for the phosphotriester hydrolysis by OPDA enzyme.

from α - and β -metals) while the substrate remains mono-coordinated to the β -metal, at a distance of 2.405 Å.

The internuclear distance between the two metal ions is 3.536 Å (vs. 3.522 Å exptl).

In the optimized complex, the substrate interacts with the close water molecule $H_2O(1)$ while a network of hydrogen-bond interactions contributes to the achievement of a better charge distribution into the active site. In particular, the observed H bond between Tyr257 and Asp301 (1.956 Å) was also the object of previous investigations which established that such interaction could justify the functional differences observed between OpdA and OPH.^[13] Moreover, Arg254 together with the water molecule $H_2O(2)$ is involved in a network of H bonds with the OH terminally bound to β -Co^{II} ion and, again, with the Asp301.

In the optimized ES complex, the distance between the nucleophilic hydroxide (Onuc) and the phosphorus atom of the incoming substrate is 3.621 Å. Thanks to this disposition in the active site, the substrate is in an ideal near-attack configuration.^[37]

Actually the subsequent TS1 involves the shortening of the P-Onuc distance which decreases from 3.621 to 1.878 Å. The analysis of the vibrational mode clearly indicates the stretching of the incoming P-Onuc bond, as a result of the nucleophilic attack of the α -Co²⁺ bound hydroxide to the phosphorus atom, and was confirmed to be a first-order saddle point with only one imaginary frequency (107_i cm⁻¹). At the same time

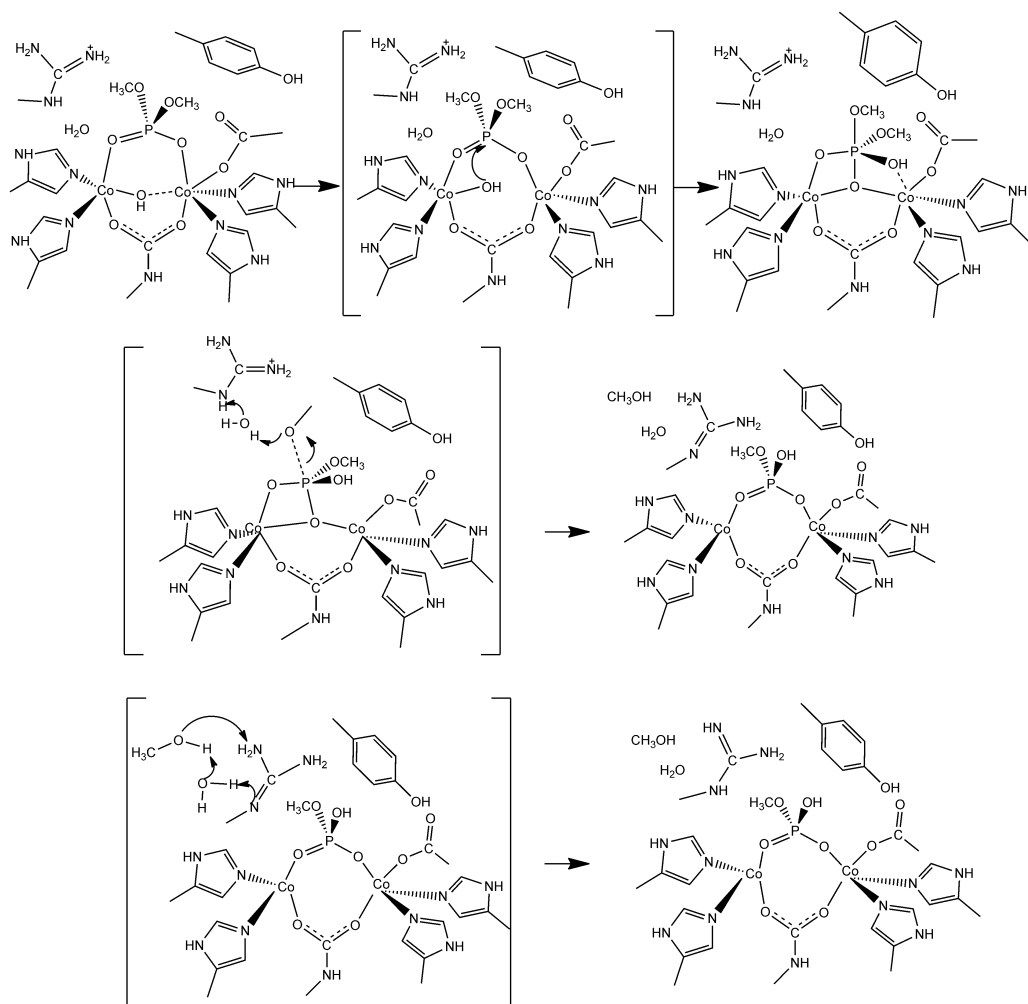
the nucleophile moves slightly away from the α -metal ion, being at a distance of 2.040 Å from it. The two water molecules are still involved in the hydrogen-bond network which introduces a significant stabilization of the system.

From an energetic point of view, the nucleophilic attack requires an energetic expense of 12.1, 8.7, 6.0, 3.7 kcal mol⁻¹ at B3LYP, B3LYP-D, M06L and MPWB1K, respectively (Figure 3).

The nucleophilic attack induces a cascade of geometric changes leading to the first enzyme-intermediate complex (INT). The phosphorus atom adopts a trigonal bipyramidal geometry in which the P-Onuc bond is completely formed, being 1.658 Å, while that between α -Co²⁺ and the nucleophile elongates until 2.434 Å. As a consequence, the pentacoordinated phosphate interacts with both metal ions in such a bicoordinated manner and it is stabilized in the active site by a network of hydrogen-bond interactions with both water molecules lying in the outer coordination sphere of metal ions.

To re-establish the coordination number of the α -metal center, which has just transferred the nucleophile to the substrate, the hydroxyl originally coordinated to the β -metal places itself halfway between the cobalt ions (2.073 and 2.004 Å from β - and α -center, respectively). The pentacoordinated intermediate lies at 16.8, 20.3, 20.8 and 15.1 kcal mol⁻¹ below the TS1, at B3LYP, B3LYP-D, MPWB1K and M06L levels, respectively.

Water molecule (1) is already in the position to act as proton-shuttle in the following transition state, TS2. Actually



Scheme 3. Proposed reaction mechanism for the phosphodiester hydrolysis by OPDA enzyme.

the hydrolysis reaction proceeds with the proton transfer from the -OH nucleophile to water (1), which in turn, transfers one of its protons to the leaving -OCH_3 group. The distance between the latter and the phosphorus atom increases from 1.740 Å in the previous INT, to 1.885 Å precluding to the release of the methanol group.

The proton shift is significantly facilitated by the intercession of the water and requires an energy cost of 10.3, 9.9, 11.9 and 12.5 kcal mol^{-1} to occur, at B3LYP, B3LYP-D, MPWB1K and M06L, respectively.

Actually, the same transition state without water assistance requires a higher energy barrier to be overcome confirming the crucial function as proton-shuttle played by water and its decisive role in the catalytic efficiency. (More details concerning the energetic profiles and optimized geometries obtained for the not-assisted mechanism can be found in the Support-

ing Information, Figure S2.) The analysis of the obtained vibrational frequency (823 cm^{-1}) clearly indicates the proton-shift processes as described above.

The substrate is tightly bicoordinated to the metals at distances of 2.230 and 2.011 Å from α - and β - Co^{II} ions.

The detachment of the protonated leaving group leads to the enzyme-product complex (EP), which is obtained exothermally with a great energy gain predicted equal to 46.9, 50.7, 51.0 and 43.5 kcal mol^{-1} , at B3LYP, B3LYP-D, MPWB1K and M06L levels. The released group lies at about 5.444 Å from the phosphorus atom and still interacts with the water molecule (1) [at a distance of 1.816 Å] and with the His230 (at a distance of 1.868 Å).

The phosphorus atom regenerates its optimal tetrahedral geometry and the resulting dimethyl-ester is coordinated equidistantly to α - and β -metal ions with distances of 2.152 and

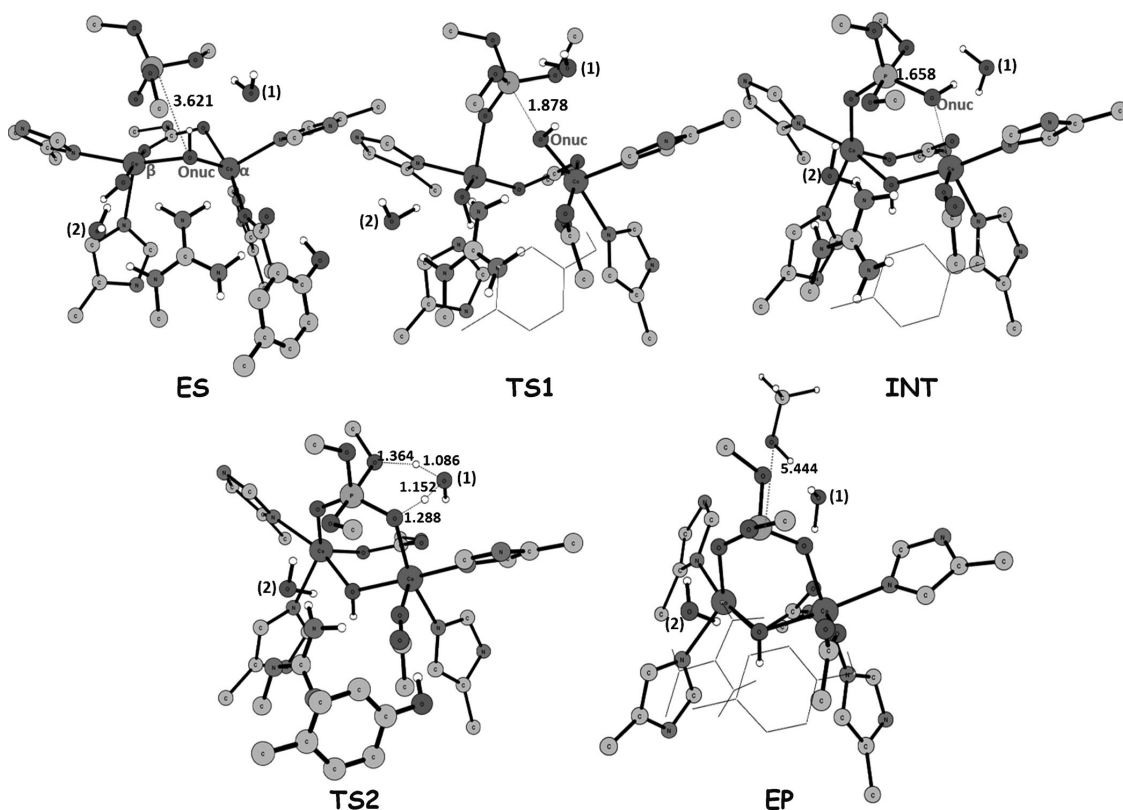


Figure 2. Optimized structures of the stationary points ES, TS1, INT, TS2, and EP for the hydrolysis of trimethyl phosphate promoted by the OpdA enzyme, with the most relevant geometrical parameters. Only the most significant hydrogen atoms are reported in the figure.

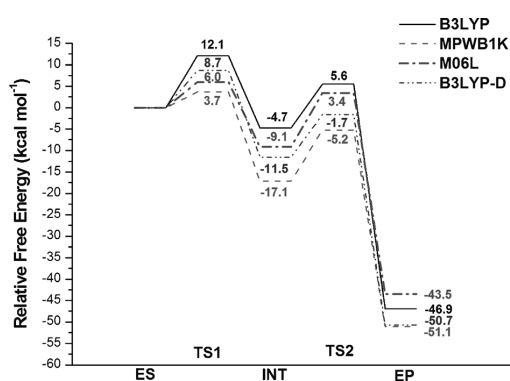


Figure 3. Relative free energy in solution for the phosphotriester hydrolysis by di-Co²⁺ OPDA enzyme.

2.109 Å, respectively. The metal-bridging hydroxide in the EP complex is located at a distance of 3.212 Å from phosphorus and it is likely to lead the next nucleophilic attack, both the diester and the possible nucleophile being optimally oriented to allow hydrolysis.

As it is well known that the cobalt(II) transition metal ion often acts as cofactor in enzymes.

The Co²⁺ ion has seven electrons in the valence shell and they are usually in a high-spin arrangement. The correct description of the electronic configuration of the d shell of binuclear centers in bimetallic enzymatic complexes is a crucial point. The evaluation of the existence of an antiferromagnetic (AFM) coupling between metals^[38] should be taken into account, since it could drastically influence the energetics of the reaction. Indeed, in a recent theoretical investigation on the Mn–Mn containing serine/threonine protein phosphatase 5,^[31a] the effect of AFM coupling was found to be determinant for the stabilization of the transition-state structure.

In the investigated system, among the two computed antiferromagnetic singlet high- and low-spin states for the enzyme–substrate complex, the first one was found to be the most stable one (Figure S3 in the Supporting Information) and just slightly less stable than the ground ferromagnetic (FM) septet spin state. For this reason we thought it would be interesting to explore the whole mechanism of the reaction considering the antiferromagnetic coupling. The gas-phase energetic path obtained for the antiferromagnetic singlet high-spin state, provided in the Supporting Information, is almost

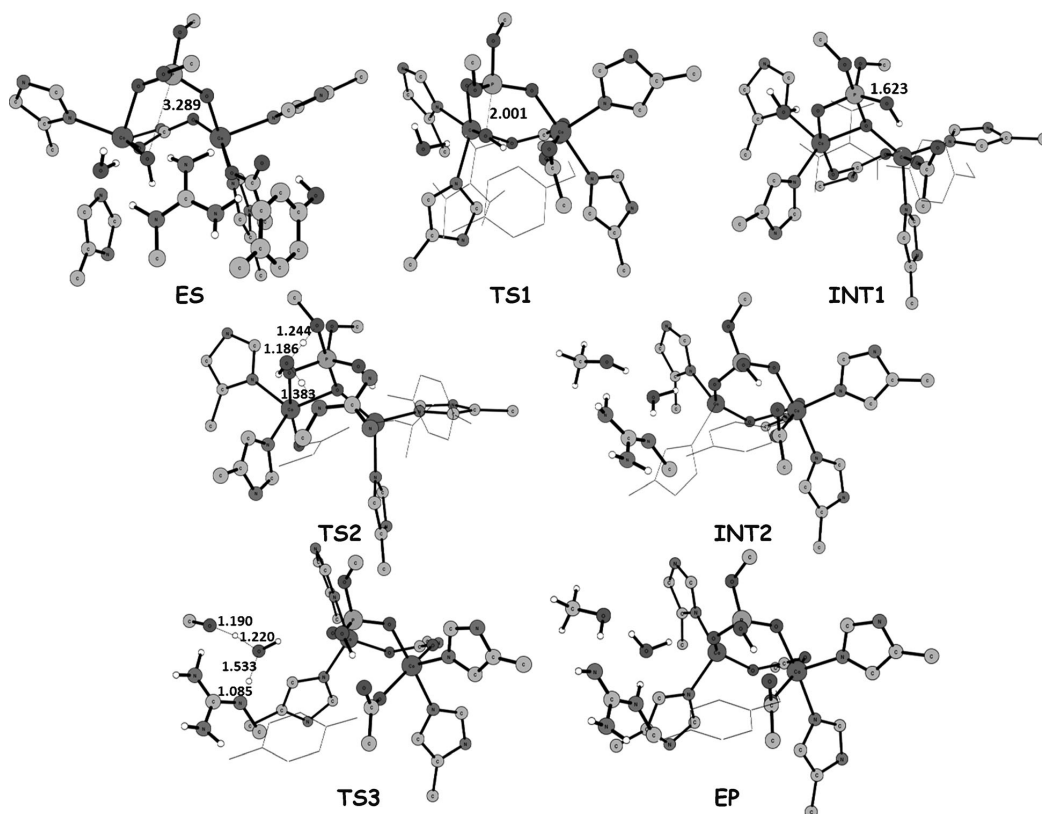


Figure 4. Optimized structures of the stationary points found for the hydrolysis of dimethyl phosphate promoted by the OpdA enzyme, with the most relevant geometrical parameters. Only the most significant hydrogen atoms are reported in the figure. Some residues are shown in thin stick.

identical to that obtained considering the FM septet spin state. This means that the two reaction pathways could be competitive, even if the septet-spin configuration is supported by experimental indications suggesting a weakly ferromagnetically coupled di-Co^{II} metal center.^[23]

Dimethyl phosphate hydrolysis

As previously mentioned, the hydrolysis reaction mechanism of the phosphodiester bond investigated in our work is depicted in Scheme 2, while structural and energetic parameters, are reported in Figure 4 and Figure 5, respectively.

The EP complex obtained in the previous process, has been used as starting point for the phosphodiester hydrolysis. The elimination of the leaving CH₃OH group and of a water molecule from that model cluster, required a re-optimization of structural and energetic parameters. The optimized enzyme–substrate system keeps some of the features of EP species even if some interesting differences can be observed. The internuclear distance between the two metal ions results to be elongated to 3.656 Å. The dimethylester is still tightly bicoordinated through its oxygen atoms to the bimetallic center

showing distances of 2.017 and 2.036 Å from α - and β -Co²⁺, respectively.

During the optimization process, the bridging hydroxide moves away from the α -metal resulting in a pseudo-monocoordination to the β -metal (3.399 and 1.974 Å, from α - and β -metal, respectively). This shift should enhance the nucleophilicity of the OH group which is well oriented to perform the attack on the phosphorus atom and lies at a distance of 3.289 Å from it. Actually, the first transition state (TS1) is characterized by the nucleophilic attack. Geometrical features indicate that the nucleophile approaches the phosphorus atom at a distance of 2.001 Å. Bond lengths between the oxygen atoms of substrate and α - and β -Co²⁺ slightly decrease (1.991 and 1.982, respectively). An imaginary frequency of -98 cm^{-1} confirms its nature as first-order saddle point. The vibrational mode well describes the formation of the P–O_{nu}c bond.

During the optimization process, the bridging hydroxide moves away from the α -metal resulting in a pseudo-monocoordination to the β -metal (3.399 and 1.974 Å, from α - and β -metal, respectively).

This shift should enhance the nucleophilicity of the OH group which is well oriented to perform the attack on the

phosphorus atom and lies at a distance of 3.289 Å from it. Actually, the first transition state (TS1) is characterized by the nucleophilic attack.

Geometrical features indicate that the nucleophile approaches the phosphorus atom at a distance of 2.001 Å. Bond lengths between the oxygen atoms of substrate and α - and β -Co²⁺ slightly decrease (1.991 and 1.982 Å, respectively).

An imaginary frequency of -98 cm^{-1} confirms its nature as first-order saddle point. The vibrational mode well describes the formation of the P-Onuc bond.

This step requires 19.3, 19.5, 16.1 and 21.7 kcal mol⁻¹ to occur, respectively, at B3LYP, B3LYP-D, MPWB1K and M06L levels of theory.

In the resulting intermediate (INT1), the phosphodiester binds in a tripodal mode to the metal ions, with one oxygen binding to the β -Co²⁺ (distance is 2.129 Å), a second oxygen bridging the two metal ions at distances of 1.945 and 1.988 Å from α - and β -metal by the displacement of the nucleophile which is now linked to the phosphorus atom (1.623 Å) and establishes a weak interaction with α -Co²⁺. Such unusual arrangement, was previously observed in the case of sweet potato purple acid phosphatase^[39] and urease.^[40] The tripodal arrangement of phosphate confirms the role of bridging -OH as the reaction nucleophile. The phosphorus atom adopts a pentacoordinated trigonal bipyramidal geometry.

The INT1 intermediate is located at 1.5 kcal mol⁻¹ below TS1 at B3LYP and MPWB1K levels of theory, while it is found to be isoenergetic with TS1 by using M06L. On the contrary, a higher stabilization of this intermediate is obtained by using B3LYP-D (5.5 kcal mol⁻¹ below TS1)

In the next stationary point (TS2) the role of the water molecule lying in the second coordination sphere is highlighted as proton-shuttle. Actually it transfers one proton to the leaving group being simultaneously protonated by the close amino acid residue Arg254.

The bond between the P atom and the involved -OCH₃ becomes longer (1.918 vs. 1.724 Å in the previous INT1) preluding to its detachment. The vibrational mode associated to the hydrogen shift mediated by water has a frequency of -848 cm^{-1} . TS2 requires energy barriers to be overcome equal to 9.1, 12.6, 11.2 and 7.7 kcal mol⁻¹ calculated at B3LYP, B3LYP-D, MPWB1K and M06L levels, respectively, leading to the subsequent stationary point INT2. Following the final departure of the methoxy group from phosphorus atom, the substrate is converted to a less damaging monoester.

The water molecule is restored and a methanol molecule is formed. It lies 5.933 Å far from the phosphorus center and is involved in a hydrogen-bond network with both Arg254 and the close water molecule.

The phosphorus atom returns to tetrahedral and is bicoordinated to its oxygen atoms to α -Co²⁺ and β -Co²⁺ metal ions with distances of 1.957 and 2.036 Å. INT2 is predicted to be a stable species although the stabilization energies given by B3LYP and the other functionals are significantly different. Indeed INT2 lies at $-15.4\text{ kcal mol}^{-1}$ below ES at B3LYP level, while it is found at -1.0 , -3.7 and $-2.7\text{ kcal mol}^{-1}$ at B3LYP-D, MPWB1K and M06L, respectively.

The enzymatic reaction proceeds to the final product via the TS3 transition state, characterized by a series of proton shifts aimed to realize a more suitable protonation state of the active site. Two routes have been explored. The first one, is the abstraction by the Arg254 imino group of one hydrogen of the water molecule which in turn rips a hydrogen from methanol and finally methanol deprotonates the -NH₂ group of the Arg residue. The vibrational frequency (-658 cm^{-1}) is clearly associated to the above-described process. In such a case, to reach the product, a barrier of 10.3, 9.4, 12.1 or 12.3 kcal mol⁻¹ must be overcome, respectively, at B3LYP, B3LYP-D, MPWB1K and M06L level. The product EP is more stable than ES by 16.0 kcal mol⁻¹ (B3LYP) while the predicted stabilization is smaller if calculated at B3LYP-D, MPWB1K and M06L (1.6, 4.1 and 2.8 kcal mol⁻¹, respectively). Despite these differences, all the XC-functionals used show the exothermic nature of the process. The product (CH₃)HPO₄ is still coordinated with its oxygen atoms to the α -Co²⁺ and β -Co²⁺ metal ions waiting for a new water molecule that moves it permanently away from active site (Figure 4). Beside this process, since the pK_a of Arg is much larger than a phosphate, we have verified the possibility of a proton transfer mediated by water from phosphate to Arg254, to have a product with protonated Arg and a deprotonated phosphate (Figure S4 in the Supporting Information). Gas phase computations show that the product formed after this proton transfer is energetically less stable by about 3 kcal mol⁻¹ than that previously obtained, and the transition state leading to it, is 1.0 kcal mol⁻¹ higher than the prior TS3 (in the Supporting Information). Nevertheless, the differences between the two explored routes decrease considering the solvent environment, suggesting that also this possibility is reasonable.

The highest energetic barrier of the proposed process, corresponding to the nucleophilic attack described by TS1, thus representing the rate-determining step for the diesterase process.

The structural analysis of the intermediates and transition states obtained from our computations gives information on the role of metals and some of the involved amino acid residues.

The metal ions help in keeping the different active-site groups together, as can be deduced by the preserved coordination of Lys169, His55, His57, Asp301, His201 and His230 to the binuclear center during the whole reaction path. In the hydrolysis of trimethylester, metals accomplish a structural and catalytic role. In the first step, the β -metal binds the substrate orienting it properly to the nucleophilic attack by the α -metal hydroxide. Subsequently they act together anchoring the substrate and allowing the release of the leaving group. In the second hydrolysis process their role seem to be mostly electrostatic since there are no changes in the coordination of the metal ions during the reaction.

Arg254 and Tyr257, belonging to the second coordination shell, act together to create a network of H bonds which from one side help the active site to achieve a better charge distribution and from the other side stabilize the system. Moreover,

Arg254 has a more determinant role in favoring the cleavage of phosphate-ester bond and in the consequent restoration of the enzyme active site.

The fundamental role of the two outer-sphere water molecules as proton carrier and their influence on the catalytic efficiency of the enzyme has been demonstrated.

Comparison between XC-functionals

The choice of the XC-functionals to achieve a realistic description of a chemical event and to accurately predict the energetic of the process has become extremely crucial, since no functional so far is accurate to predict all the properties of interest. As a consequence performing a benchmark study before embarking on a new DFT study is an absolute requirement, since none of the common functionals is good at everything, and as it is known, B3LYP, which is reasonable for several properties (especially for geometrical parameters), often fails in predicting the barrier heights and has been suggested to systematically overestimate to a large extent the weak interaction distances.^[30a] The analysis of the energetic profiles obtained in our work for dimethyl- and trimethyl phosphate hydrolysis, provides information on the performances of the four different B3LYP, B3LYP-D MPWB1K, and M06L functionals in estimating the energetic barriers involved in the processes. In the case of main enzymatic activity, the existing kinetic experimental data^[21] suggest a free energy barrier value of 12.9 kcal mol⁻¹, approximately. As shown in Figure 3, the new generation B3LYP-D, MPWB1K, and M06L, which have been suggested to be more accurate in reproducing this quantity, indicate the proton shift (TS2) necessary for leaving the -OCH₃ group as the step that requires the highest energy cost (9.8, 11.9, and 12.5 kcal mol⁻¹, at B3LYP-D, MPWB1K, and M06L, respectively). Both the last two XC-functionals reproduce with high accuracy the experimental value, while B3LYP-D predicts a smaller activation barrier.

On the contrary, the B3LYP functional proposes the nucleophilic attack as rate-determining step of the process (12.1 kcal mol⁻¹) even if just a slight difference between the first and the second free energy barrier is obtained. However, although the B3LYP proposed values, are not likely to indicate conclusively the rate-limiting step of the process, nevertheless it also reproduces accurately the value estimated from experimental data.

In the case of dimethyl phosphate hydrolysis, as it is evident from Figure 5, the most relevant differences due to the use of B3LYP, B3LYP-D, MPWB1K and M06L functionals concern the final portion of the energetic path. However, the barriers involved in the reaction are estimated in a similar way in the four sets of computations. In the absence of experimental data we cannot decide which functional gives the better indication with regard to the stability of the INT2 and the EP complex. Nevertheless, a deeper look of the structural features of the minima and the transition state characterizing the final portion of the reaction path, evidences the occurrence of long range interactions. Since MPWB1K has been previously suggested to accurately reproduce weak interaction distances^[30a] while

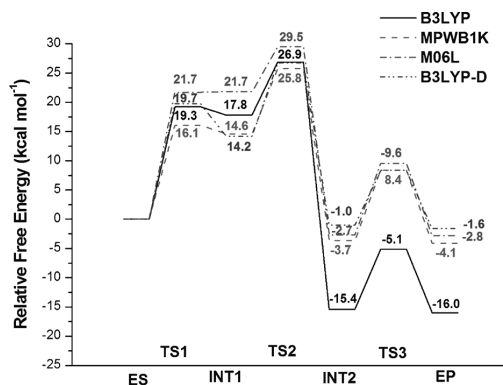


Figure 5. Relative free energy in solution for the phosphodiester hydrolysis by OPDA enzyme.

B3LYP-D and M06L includes empirical dispersion, we could hypothesize that the most reliable results are those obtained with these functionals.

Conclusion

In this study, we have reported a theoretical investigation of the reaction mechanism of an organophosphate degrading enzyme which uses two cobalt ions in its active site to carry out the catalytic conversion of a phosphotriester to phosphodiester to a phosphomonoester. The computations were performed with an active-site model consisting of 122 atoms and built up on the basis of the X-ray diffraction structure. The transition states and intermediates along the reaction pathway were located and characterized.

Our results indicate that the phosphotriester hydrolysis occurs in two steps while a further step is necessary for the next phosphodiester hydrolysis. In both cases, the beginning of the process involves the nucleophilic attack on the phosphorus atom of the substrate. This attack is led by a hydroxide terminally bound to the α -Co²⁺ ion, in the triesterase and by a bridging hydroxide placed between the two metal ions in the diesterase activities. The first step in the phosphotriester hydrolysis leads to the formation of a pentacoordinated intermediate which evolves in the more stable phosphodiester overcoming a second transition state.

Instead, the conversion of phosphodiester to phosphomonoester, after the first nucleophilic attack, requires two more steps characterized by a series of rearrangements leading to the enzyme-product complex.

Free-energy profiles have been obtained in the protein environment for the most stable ferromagnetic septet spin state. The triester hydrolysis energetic path is characterized by an activation barrier of 12.1, 9.8, 11.9, and 12.5 kcal mol⁻¹ at B3LYP, B3LYP-D, MPWB1K, and M06L levels, which reproduces with accuracy the free-energy barrier derived from the experimental K_{cat} value (≈ 12.9 kcal mol⁻¹). In the case of diester hydrolysis, the process requires a barrier to be overcome equal to 19.3, 19.5, 16.1 and 21.7 kcal mol⁻¹ at B3LYP, B3LYP-D, MPWB1K and

M06L levels, respectively. The processes are thermodynamically favored. Our study confirms that OpdA enzyme works better as triesterase as can be argued by the barrier height values in the rate-limiting steps of the catalytic processes.

With the aim to describe accurately the electronic configuration of the two cobalt ions we have also explored how an antiferromagnetic coupling could influence the energetic of the enzymatic process.

Results indicated that the antiferromagnetic singlet high-spin configuration is practically isoenergetic to that of ground ferromagnetic (FM) septet spin state. The comparison of the two reaction paths for the most stable FM and AFM couplings shows no significant differences in energy values of all minima and transition states involved in the mechanism suggesting the likelihood of competition between them.

Acknowledgements

Financial support from the Università degli Studi della Calabria-Dipartimento di Chimica e Tecnologie Chimiche (CTC) is acknowledged.

Keywords: computational chemistry · density functional calculations · enzyme catalysis · reaction mechanisms

- F. Ely, J.-L. Foo, C. J. Jackson, L. R. Gahan, D. L. Ollis, G. Schenk, *Curr. Top. Biochem. Res.* **2007**, *9*, 63–78.
- D. S. Scanlan, R. C. Reid, *Chem. Biol.* **1995**, *2*, 71.
- D. P. Dumas, S. R. Caldwell, J. R. Wild, F. M. Raushel, *J. Biol. Chem.* **1989**, *264*, 19659.
- F. M. Raushel, *Curr. Opin. Microbiol.* **2002**, *5*, 288–295.
- N. Mitic, S. J. Smith, A. Neves, L. W. Guddat, L. R. Gahan, G. Schenk, *Chem. Rev.* **2006**, *106*, 3338–3363.
- I. Horne, T. D. Sutherland, R. L. Harcourt, R. J. Russell, J. G. Oakeshott, *Appl. Environ. Microbiol.* **2002**, *68*, 3371–3376.
- D. M. Munnecke, *Appl. Environ. Microbiol.* **1976**, *32*, 7–13.
- N. Sethunathan, T. Yoshida, *Can. J. Microbiol.* **1973**, *19*, 873–875.
- E. Ghanem, Y. Li, C. Xu, F. M. Raushel, *Biochemistry* **2007**, *46*, 9032–9040.
- S. D. Aubert, Y. Li, F. M. Raushel, *Biochemistry* **2004**, *43*, 5707–5715.
- M. A. Anderson, H. Shim, F. M. Raushel, W. W. Cleland, *J. Am. Chem. Soc.* **2001**, *123*, 9246–9253.
- M. Chen-Goodspeed, M. A. Sogorb, F. Wu, S.-B. Hong, F. M. Raushel, *Biochemistry* **2001**, *40*, 1325–1331.
- G. Schenk, N. Mitic, L. R. Gahan, D. L. Ollis, R. P. McGeary, L. W. Guddat, *Acc. Chem. Res.* **2012**, *45*, 1593–1603.
- L. J. Daumann, B. Y. McCarthy, K. S. Hadler, T. Murray, L. R. Gahan, J. A. Larrabee, D. L. Ollis, G. Schenk, *Biochim. Biophys. Acta Proteins Proteomics* **2013**, *1834*, 425–432.
- C. J. Jackson, P. D. Carr, H.-K. Kim, J.-W. Liu, P. Herral, N. Mitic, G. Schenk, C. A. Smith, D. L. Ollis, *Biochem. J.* **2006**, *397*, 501–508.
- K. S. Hadler, E. A. Tanifum, S. Yip, N. Mitic, L. W. Guddat, C. J. Jackson, L. R. Gahan, K. Nguyen, P. D. Carr, D. L. Ollis, A. C. Hengge, J. A. Larrabee, G. Schenk, *J. Am. Chem. Soc.* **2008**, *130*, 14129–14138.
- F. Ely, K. S. Hadler, L. R. Gahan, L. W. Guddat, D. L. Ollis, G. Schenk, *Biochem. J.* **2010**, *432*, 565–573.
- C. Jackson, H.-K. Kim, P. D. Carr, J.-W. Liu, D. L. Ollis, *Biochim. Biophys. Acta Proteins Proteomics* **2005**, *1752*, 56–64.
- M. M. Benning, H. Shim, F. M. Raushel, H. M. Holden, *Biochemistry* **2001**, *40*, 2712–2722.
- J. Kim, P.-C. Tsai, S.-L. Chen, F. Himo, S. C. Almo, F. M. Raushel, *Biochemistry* **2008**, *47*, 9497–9504.
- C. J. Jackson, J.-W. Liu, M. L. Coote, D. L. Ollis, *Org. Biomol. Chem.* **2005**, *3*, 4343–4350.
- C. J. Jackson, J.-L. Foo, H.-K. Kim, P. D. Carr, J.-W. Liu, G. Salem, D. L. Ollis, *J. Mol. Biol.* **2008**, *375*, 1189–1196.
- F. Ely, K. S. Hadler, N. Mitic, L. R. Gahan, D. L. Ollis, N. M. Plugis, M. T. Russo, J. A. Larrabee, G. Schenk, *J. Biol. Inorg. Chem.* **2011**, *16*, 777–787.
- A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648; b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.
- Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Menucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
- T. H. Dunning, Jr., P. J. Hay, in *Modern Theoretical Chemistry* (Ed. H. F. Schaefer III), Vol. 3, Plenum, New York, **1977**, pp. 1–28.
- R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.
- a) M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, *114*, 3601–3658; b) P. E. M. Siegbahn, F. Himo, *J. Biol. Inorg. Chem.* **2009**, *14*, 643–651; c) T. Lind, P. E. M. Siegbahn, R. H. Crabtree, *J. Phys. Chem. A* **1999**, *103*, 1193–1202; d) P. E. M. Siegbahn, M. R. A. Blomberg, *Chem. Rev.* **2000**, *100*, 421–437; e) L. Noodleman, T. Lovell, W. G. Han, J. Li, F. Himo, *Chem. Rev.* **2004**, *104*, 459–508; f) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257–2261; g) P. C. Hariharan, J. A. Pople, *Mol. Phys.* **1974**, *27*, 209–214; h) M. S. Gordon, *Chem. Phys. Lett.* **1980**, *76*, 163–168.
- M. E. Alberto, M. Leopoldini, N. Russo, *Inorg. Chem.* **2011**, *50*, 3394–3403.
- a) S. F. Sousa, P. A. Fernandes, M. J. Ramos, *J. Phys. Chem. A* **2007**, *111*, 10439; b) A. J. M. Ribeiro, M. J. Ramos, P. A. Fernandes, *J. Chem. Theory Comput.* **2010**, *6*, 2281; c) S. F. Sousa, P. A. Fernandes, M. J. Ramos, *J. Phys. Chem. A* **2009**, *113*, 14231–14236.
- a) A. J. M. Ribeiro, M. E. Alberto, M. J. Ramos, P. A. Fernandes, N. Russo, *Chem. Eur. J.* **2013**, *19*, 14081–14089; b) M. E. Alberto, T. Marino, N. Russo, M. Toscano, E. Sicilia, *Phys. Chem. Chem. Phys.* **2012**, *14*, 14943–14953; c) M. E. Alberto, T. Marino, M. J. Ramos, N. Russo, *J. Chem. Theory Comput.* **2010**, *6*, 2424–2433; d) T. Marino, N. Russo, M. Toscano, *Inorg. Chem.* **2013**, *52*, 655; e) T. Marino, N. Russo, M. Toscano, *J. Am. Chem. Soc.* **2005**, *127*, 4242–4253.
- S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 6908–6918; b) Y. Zhao, D. G. Truhlar, *J. Chem. Phys.* **2006**, *125*, 194101.
- V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995.
- M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* **2003**, *24*, 669.
- F. Ely, M. M. Pedroso, L. R. Gahan, D. L. Ollis, L. W. Guddat, G. Schenk, *J. Inorg. Biochem.* **2012**, *106*, 19–22.
- S. Hur, T. C. Brulce, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12015.
- a) G. C. Dismukes, *Chem. Rev.* **1996**, *96*, 2909; b) S. V. Khangulov, P. J. Pessiki, V. V. Barynin, D. E. Ash, G. C. Dismukes, *Biochemistry* **1995**, *34*, 2015.
- G. Schenk, L. R. Gahan, L. E. Carrington, N. Mitic, M. Valizadeh, S. E. Hamilton, J. de Jersey, L. W. Guddat, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 273.
- S. Benini, W. R. Rypniewski, K. S. Wilson, S. Ciurli, S. Mangani, *J. Biol. Inorg. Chem.* **2001**, *6*, 778–790.

Received: October 10, 2014

Published online on January 12, 2015

8 Bibliografia

1. *Analysis of zinc-ligand bond lengths in metalloproteins: trends and patterns.* **Tamames B, Sousa SF, Tamames J, Fernandes PA, Ramos MJ.** 3, 2007, *Proteins*, Vol. 69, pp. 466-475.
2. *Effective tailor-made force field parameterization of the several Zn coordination environments in the puzzling FTase enzyme: opening the door to the full understanding of its elusive catalytic mechanism.* **M.J. Ramos, S.F. Sousa, P.A. Fernandes.** 1, 2007, *Theoretical Chemistry Accounts*, Vol. 117, pp. 171-181.
3. **Karlheinz Drauz, Harald Gröger, Oliver May.** *Concepts in Biocatalysis. Enzyme Catalysis in Organic Synthesis.* s.l. : Wiley, 2012, pp. 43-66.
4. **Upadhyay, Santosh K.** *Chemical Kinetics and Reaction Dynamics.* Kanpur : Springer, 2006.
5. **GM., Cooper.** <http://www.ncbi.nlm.nih.gov/>. [Online] 2000. [Citação: 14 de November de 2014.] <http://www.ncbi.nlm.nih.gov/books/NBK9921/>.
6. **Klinman, Judith, Hammes-Schiffer, Sharon.** *Dynamics in Enzyme Catalysis.* s.l. : Springer, 2013.
7. **Mueller, Michael.** *Fundamentals of Quantum Chemistry.* New York : Kluwer, 2002.
8. *Inhomogeneous Electron Gas.* **Hohenberg, P. and W. Kohn.** 3B, 1964, *Physical Review*, Vol. 136, pp. 864-&.
9. **Wolfram Koch, Max C. Holthausen.** *A Chemist's Guide to Density Functional Theory.* Weinheim : Wiley, 2001.
10. **Eschrig, Helmut.** *The fundamentals of DFT.* s.l. : Gutenbergplatz, 2003.
11. *Note on exchange phenomena in the Thomas atom.* **P.A.M., Dirac.** 1930, *Proceedings of the Cambridge Philosophical Society*, Vol. 26, pp. 376-385.
12. *General performance of density functionals.* **Ramos, M.J., S.F. Sousa, and P.A. Fernandes.** 42, 2007, *Journal of Physical Chemistry A*, Vol. 111, pp. 10439-10452.
13. *A DFT and ab Initio Benchmarking Study of Metal-Alkane Interactions and the Activation of Carbon-Hydrogen Bonds.* **Flener-Lovitt C, Woon DE, Dunning TH Jr, Girolami GS.** 4, 2009, *The Journal of Physical Chemistry A*, Vol. 114, pp. 1843-1851.
14. *"Mindless" DFT Benchmarking.* **Korth, M. and S. Grimme.** 4, 2005, *Journal of Chemical Theory and Computation*, Vol. 5, pp. 993-1003.
15. *Density functional thermochemistry. I. The effect of the exchange only gradient correction.* **Becke, A.D.** 1992, *The Journal of chemical physics*, Vol. 96, pp. 2155-&.
16. *Generalized gradient approximation made simple.* **J. P. Perdew, K. Burke, and M. Ernzerhof.** 1996, *Phys. Rev. Lett.*, Vol. 77, pp. 3865-68.
17. *Errata: Generalized gradient approximation made simple.* **J. P. Perdew, K. Burke, and M. Ernzerhof.** 1997, *Phys. Rev. Lett.*, Vol. 78, p. 1396.
18. *Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average than B3LYP for Ground States.* **Truhlar, Y. Zhao and D. G.** 2006, *J. Phys. Chem. A*, Vol. 110, pp. 13126-13130.
19. *Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids.* **J. M. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria.** 2003, *Phys. Rev. Lett.*, Vol. 91, p. 146401.
20. *Density functional thermochemistry. II. The effect of the Perdew-Wang generalized gradient correlation correction.* **Becke, A.D.** 1992, *The Journal of chemical physics*, Vol. 97, pp. 9173-&.
21. *Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing.* **Becke, A.D.** 3, 1996, *Journal of Chemical Physics*, Vol. 104, pp. 1040-1046.
22. *Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals.* **Becke, A.D.** 1997, *The Journal of chemical physics*, Vol. 107, pp. 8554-8562.
23. *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis.* **Vosko, S.H., Wilk, L. Nusair, M.** 1980, *Can. J. Phys.*, Vol. 58, pp. 1200-&.
24. *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.* **Lee, C., Yang, W. Parr, R.G.** 1988, *Phys. Rev B*, Vol. 37, p. 785.

25. *Self-Consistent Molecular Orbital Methods. 1. Use of Gaussian expansions of Slater-type atomic orbitals.* **W. J. Hehre, R. F. Stewart, and J. A. Pople.** 1969, *J. Chem. Phys.*, Vol. 51, pp. 2657-64.
26. *Self-Consistent Molecular Orbital Methods. 17. Geometries and binding energies of second-row molecules. A comparison of three basis sets.* **J. B. Collins, P. v. R. Schleyer, J. S. Binkley, and J. A. Pople.** 1976, *J. Chem. Phys.*, Vols. 5142-51, p. 64.
27. *A Proper Account of Core-polarization with Pseudopotentials - Single Valence-Electron Alkali Compounds.* **P. Fuentealba, H. Preuss, H. Stoll, L. v. Szentpály,** 1982, *Chem. Phys. B*, Vol. 89, pp. 418-22.
28. *An all atom force field for simulations of proteins and nucleic acids.* **Weiner, S.J.** 2, 1986, *Journal of Computational Chemistry*, Vol. 7, pp. 230-252.
29. *A New Force-Field for Molecular Mechanical Simulation of Nucleic-Acids and Proteins.* **Weiner, S.J.** 3, 1984, *Journal of the American Chemical Society*, Vol. 106, pp. 765-784.
30. *A second generation force field for the simulation of proteins, nucleic acids, and organic molecules.* **Cornell, W.D.** 19, 1995, *Journal of the American Chemical Society*, Vol. 117, pp. 5179-97.
31. *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations.* **Duan, Y.** 16, 2003, *Journal of Computational Chemistry*, Vol. 24, pp. 1999-2012.
32. **D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wol.** *AMBER12.* San Francisco : University of California, 2012.
33. *Automatic atom type and bond type perception in molecular mechanical calculations.* **Wang, J., Wang, W., Kollman P. A. e Case, D. A.** 2006, *Journal of Molecular Graphics and Modelling* , Vol. 25, p. 247260.
34. *Development and testing of a general AMBER force field.* **Wang, J., Wolf, R. M., et al.** 2005, *Journal of Computational Chemistry*, Vol. 24, pp. 1157-1174.
35. *Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development.* **R.J. Woods, R.A. Dwek, C.J. Edge, B. Fraser-Reid.** 1995, *J. Phys. Chem*, Vol. 99, pp. 3832-46.
36. *A New ONIOM Implementation in Gaussian 98. 1. The Calculation of Energies, Gradients and Vibrational Frequencies and Electric Field Derivatives.* **S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. J. Frisch.** 1999, *J. Mol. Struct. (Theochem)*, Vol. 462, pp. 1-21.
37. *Crystal structures of Fe²⁺ dioxygenase superoxo, alkylperoxo, and bound product intermediates.* **E. G. Kovaleva, Lipscomb J. D.** 5823, 2007, *Science*, Vol. 316, pp. 453-457.
38. *The Ins and Outs of Ring-Cleaving dioxygenases.* **Vaillancourt, Frederic H., Bolin, Jeffrey T. e Eltis, Lindsay D.** 4, 2006, *Critical Reviews in Biochemistry and Molecular Biology* , Vol. 41, pp. 241-267.
39. *Ring-Cleaving Dioxygenases with a Cupin Fold.* **Fetzner, Susanne.** 8, 2012, *Applied and Environmental Microbiology* , Vol. 78, pp. 2505-2514.
40. *Microbial degradation of aromatic compounds — from one strategy to four.* 11, 2011, *Nature Reviews Microbiology* , Vol. 9, pp. 803-816.
41. *Mechanism of extradiol aromatic ring-cleaving dioxygenases.* **Lipscomb, John D.** 6, 2008, *Current Opinion in Structural Biology* , Vol. 18, pp. 644-49.
42. *Dioxygen Activation at Mononuclear Nonheme Iron Active Sites: Enzymes, Models, and Intermediates.* **Costas, M, Mehn, MP e Jensen, MP.** 2, 2004, *Chemical Reviews* , Vol. 104, pp. 939-986.
43. *Aromatic Ring Cleavage of 2-Amino-4-tert-butylphenol by a Nonheme Iron(II) Complex: Functional Model of 2-Aminophenol Dioxygenases.* **Chakraborty, Biswarup e Paine, Tapan Kanti.** 3, 2013, *Angewandte Chemie International Edition* , Vol. 52, pp. 920-924.
44. *On the reliability of semi-empirical pseudopotentials - simulation of Hartree-Fock and Dirac-Fock results.* **P. Fuentealba, H. Stoll, L. v. Szentpály, P. Schwerdtfeger, and H. Preuss.** 1983, *J. Phys. B*, Vol. 16, pp. 323-28.
45. **U. Wedig, M. Dolg, H. Stoll, and H. Preuss.** *Quantum Chemistry: The Challenge of Transition Metals and Coordination Chemistry.* s.l. : Veillard, Reidel, and Dordrecht, 1986.
46. *The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic*

testing of four M06-class functionals and 12 other. **Truhlar, Y. Zhao and D. G.** 2008, *Theor. Chem. Acc*, Vol. 120, pp. 215-41.

47. *Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein.* **Kyndt, JA, Meyer, TE e Cusanovich, MA.** 2005, *FEBS Lett*, Vol. 512, pp. 240-244.

48. *Heterologous Production of Halorhodospira halophila Holo-Photoactive Yellow Protein through Tandem Expression of the Postulated Biosynthetic Genes.* **Kyndt, JA, Vanrobaeys, F e Fitch, JC.** 2003, *Biochemistry*, Vol. 42, pp. 965-970.

49. *Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete Saccharothrix espanaensis.* **Berner, M, Krug, D e Bihlmaier, C.** 2006, *J. Bacteriol*, Vol. 188, pp. 2666-2673 .

50. *Exploring Recombinant Flavonoid Biosynthesis in Metabolically Engineered Escherichia coli.* **Watts, KT, Lee, PC e Schmidt-Dannert, C.** 2004, *ChemBioChem*, Vol. 5, pp. 500-507.

51. *Using Unnatural Protein Fusions to Engineer Resveratrol Biosynthesis in Yeast and Mammalian Cells.* **Zhang, Yansheng, Li, Song-Zhe e Li, Jia.** 2006, *J. Am. Chem. Soc.* , Vol. 128, pp. 13030-13031.

52. *Improved production of p-hydroxycinnamic acid from tyrosine using a novel thermostable phenylalanine/tyrosine ammonia lyase enzyme.* **Xue, Z., et al.** 2007, *Enzyme Microb. Technol.* , Vol. 42, pp. 58-64.

53. *Functional expression in Escherichia coli of the tyrosine-inducible tyrosine ammonia-lyase enzyme from yeast Trichosporon cutaneum for production of p-hydroxycinnamic acid.* **Vannelli, T., et al.** 2007, *Enzyme Microb. Technol*, Vol. 41, pp. 413-422.

54. *Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. .* **Schwede, T. F., Rétey, J. e Schulz, G. E.** 1999, *Biochem*, Vol. 13, pp. 1327-38.

55. *Structural determinants and modulation of substrate specificity in phenylalanine-tyrosine ammonia-lyases.* **Louie, G. V., et al.** 2006, *Chem. Biol.*, Vol. 13, pp. 1327-38.

56. *Autocatalytic Peptide Cyclization during Chain Folding of Histidine Ammonia-Lyase.* **Baedeker, M. e Schulz, G. E.** 2002, *Structure*, Vol. 10, pp. 61-67.

57. *Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family.* **Watts, K. T., et al.** 2006, *Chem. Biol.* , Vol. 13, pp. 1317-26.

58. *Contributions of conserved serine and tyrosine residues to catalysis, ligand binding, and cofactor processing in the active site of tyrosine ammonia lyase.* **Schroeder, A. C., et al.** 2008, *Phytochemistry*, Vol. 69, pp. 1496-1506.

59. *Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase.* **Ritter, H. e Schulz, G. E.** 2004, *Plant cell* , Vol. 16, pp. 3426-36.

60. *The Role of Dehydroalanine in Catalysis by Histidine Ammonia Lyase.* **Langer, M., Pauling, A. e Rétey, J.** 1995, *Angewandte Chemie International Edition in English* , Vol. 34, pp. 1464-1465.

61. *Use of nitrogen-15 and deuterium isotope effects to determine the chemical mechanism of phenylalanine ammonia-lyase.* **Hermes, J. D., Weiss, P. M. e Cleland, W.** 1984, *Biochemistry* , Vol. 24, pp. 2959-2967.

62. *Mechanism of the tyrosine ammonia lyase reaction-tandem nucleophilic and electrophilic enhancement by a proton transfer.* **Pilbák, S., Farkas, Ö e Poppe, L.** 2012, *Chemistry*, Vol. 18, pp. 7793-7802.

63. *The PyMOL Molecular Graphics System.* **Schrödinger, L. L. C.** 2010.

64. *GaussView Version 5.* **Dennington II, R., Keith, T. e Millam, J.** 2009, Semichem Inc.

65. *Gaussian 09.* **risch, M. J., et al.** 2009, Gaussian Inc Wallingford CT.

66. *Self-Consistent Equations Including Exchange and Correlation Effects.* **Kohn, W. and L.J. Sham.** 4A, 1965, *Physical Review*, Vol. 140, pp. 1133-&.

67. *Density-functional exchange-energy approximation with correct asymptotic behavior.* **Becke, A.D.** 1988, *Phys. Rev. A*, Vol. 38, p. 3098.