DOCTORAL THESIS

European Joint Doctorate on Theoretical Chemistry and Computational Modelling

"Theoretical Studies of the Catalytic Mechanism of the Dihydroxyacetone Kinase"

ISABEL BORDES PASTOR

Castelló, September 2017

Escola Superior de Tecnologia i Ciències Experimentals

Departament de Química Física i Analítica

Àrea de Química Física

Vicent Moliner Ibáñez, profesor Catedrático del Departament de Química Física i Analítica de la Universitat Jaume I de Castellón y **Raquel Castillo Solsona**, profesora Titular del Departament de Química Física i Analítica de la Universitat Jaume I de Castellón,

CERTIFICAN:

Que el trabajo con título "Theoretical Studies of the Catalytic Mechanism of the Dihydroxyacetone Kinase" ha sido realizado por Isabel Bordes Pastor bajo nuestra dirección para optar al grado de Doctora en Química.

Así, autorizan la presentación de este trabajo a efectos de seguir los trámites correspondientes de la Universitat Jaume I de Castellón.

Castellón de la Plana, a 12 de Junio de 2017.

Dr. Vicent Moliner Ibáñez **Dra. Raquel Castillo Solsona**

Director de Tesis Directora de Tesis

A Javi

A mis padres y hermanos

A mi familia

A mis amigos

Acknowledgements/Agradecimientos

Durante estos años como estudiante de doctorado, además de especializarme, he tenido la suerte de vivir momentos que serán inolvidables y de poder compartirlos con gente que ya formaban parte de mi vida o que han entrado en ella gracias a esta etapa como investigadora. Es cierto que también ha habido momentos de incertidumbre, pero aun así, es inevitable quedarme con todas las experiencias buenas de las que he disfrutado durante estos años. Así pues, me gustaría mostrar mi agradecimiento a todas aquellas personas que me han acompañado de una manera u otra durante este viaje y me disculpo con antelación si me olvido de alguien.

En primer lugar, quiero dar las gracias a mis directores de tesis, Vicent Moliner y Raquel Castillo por haberme iniciado en el campo de la química computacional, por los conocimientos que me han transmitido y por guiarme durante estos años en la realización de este trabajo de investigación.

Además, me gustaría extender mi agradecimiento a todos aquellos que también me han ayudado durante estos años como estudiante de doctorado, tanto en el ámbito de trabajo como en el personal.

A Javier Ruiz, por la ayuda que me ha ofrecido con sus ideas y en materia de programación, por su apoyo constante y el optimismo que siempre me ha transmitido.

Al Dr. Aurélien de la Lande, por haberme dado la oportunidad de trabajar con él en dos ocasiones en la Universidad de Orsay, en Francia. Por sus aportaciones tanto en el trabajo como en otras cuestiones durante mi estancia allí. A mis compañeros de laboratorio en Orsay, especialmente a Natacha, Aurelie, Cristoff y David porque hicieron que me adaptara sin ningún tipo de problema.

A Kasia, por el tiempo que trabajamos juntas y por haberme dado la oportunidad de publicar un artículo con ella.

A mis compañeros de trabajo, Silvia, Lourdes, Maite, Thiago, Marisa, Kemel y Eriquin por su amistad y por todo su cariño. Con ellos he compartido el día a día en el trabajo y muchos momentos fuera de él en comidas, cenas y bailes en las noches de fiesta. Quiero acordarme también de los nuevos integrantes del grupo Daria, Miquel y Natalia con los que, de igual manera, he compartido muy buenos momentos en poco tiempo. A los profesores del departamento de Química Física y Analítica: Sixte Safont, Juan Andrés Bort, Rosa Llusar, Armando Beltrán, Sergio Martí y Mónica Oliva, por sus aportaciones durante todos estos años.

También quiero acordarme de los compañeros que ya no están en el laboratorio, a los que tengo mucho cariño: Conchín, Nacho y Pato.
Agradezco también a la secretaria del departamento, Merche, toda la ayuda que me ha ofrecido en los trámites administrativos.
De igual manera, quisiera expresar mi agradecimiento a los integrantes del grupo "*Materiales Moleculares*", tanto a los que están en el departamento a día de hoy como a los que estuvieron, por su cariño, su apoyo y por los momentos divertidos que hemos compartido. Ellos son Iván, Elena, Emma, Carmina y especialmente Eva y Tomás.
Me gustaría recordar a los amigos que hice durante el tiempo que estuve en Paris, con los que sigo manteniendo el contacto aunque pasen los años: Ana, Gaizka, Alicia y Breogan. Pasamos grandes momentos juntos y también gracias a ellos, tengo un bonito recuerdo de mi estancia allí.
Asimismo, estoy agradecida a mis compañeros del máster a los que guardo mucho cariño, especialmente a Marcos, con el que compartí mi estancia en Paris y al que incluyo en el grupo de amigos antes nombrado y a Lucia y David con los que me divertí muchísimo.
No quisiera olvidarme de los integrantes del grupo "*Modeling Chemical Processes in Biological Environments*" de Valencia: al profesor y director del grupo Iñaki Tuñón, por sus clases en el máster y por el apoyo recibido de él, a Rafa, con el que compartí el máster y la estancia en Paris con el resto de amigos y a Kirill, también compañero de máster.
En cuanto a mi vida extra laboral, me gustaría dar las gracias a mis amigos de Onda y de Altura, especialmente a mis amigas Melisa, Paula y Almudena por todo el apoyo que me han ofrecido y por estar siempre ahí, tanto en los buenos como en los malos momentos.
De la misma manera, agradezco a toda mi familia, principalmente a mis padres Pura y Andrés, a mis hermanos Sara, Cristina, Raquel, Jesús, Gloria y Javier y a mi prima Maria José, por estar en todo momento de manera incondicional y por todo el afecto que me dan.
Y como no, estoy muy agradecida a mi marido Javi, la persona que siempre está a mi lado. Por estar siempre animándome y transmitiéndome tranquilidad en los momentos de estrés, por su comprensión, su bondad, su paciencia y su apoyo absoluto.
Finalmente, agradezco a las fundaciones y organismos que han financiado mi investigación:

- A la Universitat Jaume I por las becas de investigación que se me han concedido a cargo de varios proyectos dirigidos por Sixte, Javier Ruiz y Armando Beltrán.
- Al antiguo Ministerio de Educación y al programa de ayudas *Drac*, por las becas de movilidad.
- Al programa *High Performance Computing-Europa2* (HPC-Europa2), por la beca de movilidad y las horas de computación proporcionadas para mi estancia en Francia.
-Al Ministerio de Ciencias e Innovación, por el contrato a cargo de un proyecto cuyo investigador principal era Vicent Moliner.
- A la Universitat Jaume I que a través del *Programa Grups d'Investigació d'Alt Rendiment*, me ha permitido disfrutar de varios contratos durante el doctorado.
- A la Generalitat Valenciana que a través del proyecto de investigación PROMETEO, me ha permitido disfrutar de varios contratos durante el doctorado.
- Al Ministerio de Economía y Competitividad por un contrato disfrutado a cargo de un proyecto dirigido por Maite Roca.

Abstract (English version)

Kinases are enzymes that catalyse phosphorylation reactions which are considered greatly important due to its participation in processes taking place in the cells of living beings. Kinases are related with a variety of illnesses including cancer, therefore, a significant part of the research in these systems are focused in the discovery of kinase inhibitors drugs. Moreover, these enzymes have a relevant feature; the sharing of a conserved active site which means that the same drug can affect different kinases. This characteristic also implies that the catalytic mechanism must be similar in all the family.

The present PhD Thesis is focused in the study of the reaction mechanisms of the dihydroxyacetone kinases (DHAKs) based on hybrid quantum mechanics/molecular mechanics (QM/MM) methodology. DHAKs catalyse the phosphoryl transfer from adenosine triphosphate (ATP) to dihydroxyacetone (Dha) generating adenosine diphosphate (ADP) and Dha phosphate (Dha-P). The first study reported in the present PhD Thesis is the molecular mechanism of this reaction in aqueous solution employing different semiempirical Hamiltonians to describe the QM sub-set of atoms. *Spline* corrections on the resulting free energy surfaces (FES) and DFT/MM potential energy surfaces (PES) were additionally conducted. Thus, a comparative structural, mechanistic and energetic study of different mechanisms (concerted and stepwise mechanisms) were performed employing different levels of theory. The calculations showed a strong dependency with the employed level of theory, which is a key result to take into account for the next studies. Then, the reaction was studied in the DHAK of the *Escherichia coli* (*E. coli*) organism using a computational strategy based in the previous results in aqueous solution. Two mechanisms were explored in detail: a *substrate-assisted mechanism* (being ATP the substrate) and an *asp-assisted mechanism*. Results determined as a more energetically favorable mechanism the process assisted by the substrate.

In the last part of the present PhD Thesis, and in collaboration with the experimental group of Prof. Eduardo García-Junceda (Institute of General Organic Chemistry of CSIC, Madrid), the efforts were focused on tuning the phosphoryl donor specificity of DHAK of the *Citrobacter Freundii* (*C. freundii*) organism from ATP to an inorganic polyphosphate (poly-P) based on protein engineering techniques in combination with computational methods. The directed evolution program conducted by the group of E. García-Junceda produced a mutant with remarkable poly-P activity. Therefore, in an attempt

to elucidate the origin of this activity in the modified enzyme, the poly-P binding step in the wild type protein and in this Glu526Lys active mutant was analyzed by means of computational techniques. Results suggested that this mutation favored a more adequate position of the poly-P in the active site for the reaction to take place. Finally, in order to analyse whether the activity of the Glu526Lys mutant can be also attributed to an effect in the chemical step, the phosphoryl transfer reaction from poly-P to Dha in the wild type and in this active mutant form of DHAK were explored. Results suggested negligible effects of the mutation on the chemical step, thus confirming our previous results.

Abstract (Spanish version)

Las quinasas son enzimas que catalizan las reacciones de fosforilación, consideradas de gran importancia debido a que participan en procesos que tienen lugar en las células de los seres vivos. Las quinasas están relacionadas con varios tipos de enfermedades, incluyendo el cáncer, por lo que una parte importante de la investigación en estos sistemas se está centrando en el descubrimiento de fármacos que actúen como inhibidores de estas enzimas. Además, poseen la característica de que comparten un centro activo prácticamente similar en todas ellas, por lo que un mismo fármaco puede actuar en diferentes quinasas. Esta cualidad también conlleva que el mecanismo catalítico sea similar en toda la familia.

La presente Tesis Doctoral se ha centrado en el estudio de los mecanismos de reacción de las dihidroxiacetona quinasas (DHAK) empleando la metodología híbrida de mecánica cuántica/mecánica molecular (QM/MM). Las DHAKs catalizan la transferencia del grupo fosfato desde el adenosín trifosfato (ATP) a la dihidroxiacetona (Dha) generando adenosín difosfato (ADP) y Dha fosfato (Dha-P). El primer estudio presentado en esta Tesis Doctoral es el mecanismo molecular de esta reacción en disolución acuosa empleando diferentes Hamiltonianos semiempíricos para describir el subconjunto de átomos QM. Se han realizado correcciones *spline* sobre las superficies de energía libre (FES) resultantes y también se han elaborado superficies de energía potencial (PES) al nivel DFT/MM. Así pues, se llevó a cabo un estudio comparativo estructural, mecanístico y energético de los diferentes mecanismos (concertado y por pasos) empleando los diferentes niveles de teoría. Los cálculos mostraron una fuerte dependencia con el nivel de teoría empleado, resultado clave a tener en cuenta en estudios posteriores. A continuación, se estudió la reacción en la DHAK del organismo *Escherichia coli* (*E. coli*) empleando una estrategia computacional basada en los resultados obtenidos en disolución acuosa. Se exploraron dos mecanismos: el mecanismo asistido por el sustrato (donde el ATP es el sustrato) y el mecanismo asistido por el residuo Asp. Los resultados obtenidos determinaron como el mecanismo de reacción energéticamente más favorable el proceso asistido por el sustrato.

En la última parte de la presente Tesis Doctoral y en colaboración con el grupo experimental del Prof. Eduardo García-Junceda (Instituto de Química Orgánica General del CSIC, en Madrid), el trabajo se centró en modificar la especificidad de la DHAK del organismo *Citrobacter Freundii* (*C. freundii*) desde el ATP al polifosfato inorgánico (poly-P) como dador de grupos fosfato,

por medio de técnicas de ingeniería de proteínas en combinación con métodos computacionales. El programa de evolución dirigido llevado a cabo por el grupo de E. García-Junceda produjo una mutante que presentaba actividad notable con el poly-P. En un intento de esclarecer el origen de esta actividad en la enzima modificada, se analizó la etapa de unión del poly-P en la proteina nativa y en la enzima con la mutación Glu526Lys, por medio de técnicas computacionales. Los resultados sugirieron que esta mutación favorecía una posición más adecuada del poly-P en el centro activo, facilitando así que la reacción tuviera lugar. Finalmente, para analizar si la actividad observada en la enzima mutada podía atribuirse también a un efecto en la etapa química, se exploró la reacción de transferencia de fosfato desde el poly-P hasta la Dha en las dos enzimas. Los resultados sugirieron efectos menores de la mutación en el paso químico, confirmando así los resultados obtenidos en nuestro estudio previo.

List of publications

The present PhD Thesis is based on the following publications:

- "A computational study of the phosphoryl transfer reaction between ATP and Dha in aqueous solution". Isabel Bordes, José Javier Ruiz-Pernía, Raquel Castillo and Vicent Moliner. *Organic & Biomolecular Chemistry* **2015**, 13, 10179-10190.
- "A theoretical study of the phosphoryl transfer reaction from ATP to Dha catalysed by DhaK from Escherichia coli". Isabel Bordes, Raquel Castillo and Vicent Moliner. ASAP in *Journal of Physical Chemistry B*. DOI: 10.1021/acs.jpcb.7b04862
- "Tuning the phosphoryl donor specificity of dihydroxyacetone kinase from ATP to inorganic polyphosphate. An insight from computational studies". Israel Sánchez-Moreno, Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía, Vicent Moliner and Eduardo García-Junceda. *International Journal of Molecular Sciences* **2015**, 16, 27835-27849
- "A Computational Study of the Phosphoryl Donor Activity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate". Isabel Bordes, Raquel Castillo and Vicent Moliner. Submitted for publication to *International Journal of Quantum Chemistry*.

Other publications:

- "Is Promiscuous CALB a Good Scaffold for Designing New Epoxidases?" Isabel Bordes, José Recatalá, Katarzyna Swiderek and Vicent Moliner. Molecules **2015**, 20, 17789-17806.

Contribution Report:

The author's contribution to all the papers was to perform calculations, to prepare figures and tables and to collaborate in analyzing the data and writing the manuscript.

List of abbreviations

Dihydroxyacetone kinase	DHAK
Dihydroxyacetone	Dha
Adenosine triphosphate	ATP
Dihydroxyacetone phosphate	Dha-P
Escherichia coli	E. coli
Citrobacter freundii	C. freundii
Adenosine diphosphate	ADP
Phosphoenol pyruvate carbohydrate	PEP
Phosphotransferase system	PTS
cAMP-dependent protein kinase	PKA
Cyclin-dependent kinase	CDK2
Insulin receptor kinase	IRK
Deoxyuridine triphosphate	dUTP
Deoxyuridine triphosphate nucleotidohydrolase	dUTPase
Quantum mechanics/molecular mechanics	QM/MM
Molecular dynamics	MD
Potential Energy Surface	PES
Free Energy Surface	FES
Potential of mean force	PMF
Density Functional Theory	DFT
Parametric method number 3	PM3
Parametric method number 6	PM6
Hartree-Fock	HF

Becke three-parameter Lee-Yang-Parr	B3LYP
Austin Model 1	AM1
The modified semiempirical Hamiltonian AM1/d-Prot	AM1d
Moller-Plesset of order 2	MP2
Reactants	R
Transition state	TS
Intermediate	I
Products	P
Optimized potential for liquid simulations	OPLS
Transferable intermolecular potential 3P	TIP3P
Assisted Model Building with Energy Refinement	AMBER
Chemistry at Harvard macromolecular mechanics	CHARMM
Root Mean Square Deviation	RMSD
Root Mean Square Fluctuation	RMSF
Kinetic Isotope Effects	KIE

Contents

1. Introduction
(English version)

Proteins are the most abundant organic molecules in cells. They exist in every part of each living being since they are fundamental in the structure and cellular function. Enzymes are proteins specialized in the catalysis of biological reactions and are among the most remarkable of the known biomolecules due to their extraordinary specificity and catalytic power.[1] In addition, enzymes are capable of functioning in very mild conditions of temperature and pH. One of the fundamental requirements for live is that the organism must be able to catalyse chemical reactions in an efficient and selective way. Thus, enzymes are responsible of catalysing reactions corresponding to the break down of nutrient molecules, synthesis of biological macromolecules from simple precursors and conservation and transformation of chemical energy.[2,3] Research concerning enzymes has a huge relevance considering that their anomalous degree of presence and activity in living beings may be related with diseases. Depending on the conditions of a particular disease, a deficiency or total absence of one or more enzymes could be the cause (e.g. inheritable genetic disorders), however, an excessive activity of an enzyme may be the reason as well. The diagnosis can be obtained through measurements of the activities of enzymes in blood plasma, erythrocytes or tissues specimens.[3]

Numerous research works have been done to understand the origin of the catalytic power of the enzymes.[4-9] A deep comprehension of the enzymatic mechanisms, that is, understanding the way in which enzymes catalyse chemical reactions is essential to analyse biochemical processes and can contribute to produce new medicines and catalysts.[10] To achieve this comprehension, one elementary point to consider is the comparison of the enzyme catalysed reaction with the uncatalysed equivalent one (the reaction in aqueous solution).[10] It should be pointed out that the chemical reactivity of molecules is mainly determined by relative energies of reactants, transition state/s, possible intermediate/s and products. These critical structures are along a reaction path that connects reactants to products through the minimum energy.

Computational chemistry has become a very useful tool to study the enzymatic reactivity.[11] It is capable of providing energetic and structural information about transition state and intermediate structures and it can make perceivable essential biochemical processes such as protein folding, membrane transport, drug binding and conformational changes crucial to the purpose of proteins. All of these are processes that cannot directly observed through experimental techniques.[10,12] Thus, enzymatic

simulations contribute in the comprehension of the extraordinary rate enhancement in enzymes and are able to elucidate different reaction mechanisms.[6] The reaction free energy barriers obtained in the simulations can be compared with experimental measurements assisting in establishing the mechanism behind the experimental kinetics. In addition, improvements in methods to computationally study the reactions together with the development of new highly parallel computer architectures allow increasingly better to reproduce the chemical reactivity.[13] Different methods can be used for modelling enzymatic reactions and the selection of the proper one is an important task.

In this PhD Thesis a study of the catalytic process of kinases, concretely in the dihydroxyacetone kinases (DHAKs), is presented. These proteins catalyse the phosphoryl transfer reaction from adenosine triphosphate (ATP) to dihydroxyacetone (Dha) generating dihydroxyacetone phosphate (Dha-P) and adenosine diphosphate (ADP). Kinases have many relevant properties and applications which make their study interesting. A representative DHAK from each of the two main existent classes has been chosen: DHAK from the organism *Escherichia coli* (*E. coli*) and DHAK from the organism *Citrobacter Freundii* (*C. freundii*). The methodology employed in the studies presented in this PhD Thesis is based on the use of hybrid Quantum Mechanical/Molecular Mechanical (QM/MM) potentials.

The first contribution of this PhD Thesis is the study of the phosphoryl transfer reaction from ATP to Dha in aqueous solution (Contribution I in Chapter 7). It was carried out employing different semiempirical Hamiltonians (AM1d, PM3 and PM6) for describing the QM region while the MM subsystem was treated within the OPLS (to treat part of the ATP) and TIP3P (for the solvent water molecules) classical force fields. Free energy surfaces (FES) from their previous potential energy surfaces (PES) were explored for the concerted and the stepwise mechanisms, as well as considering a mechanism assisted by the solvent water molecule. Therefore, a comparative structural, mechanistic and energetic study of the proposed mechanisms was performed using the different semiempirical Hamiltonians. Bearing in mind the limitations of semiempirical methods for describing this kind of reactions, *spline* corrections on the FESs were conducted at B3LYP density functional theoretical level with the 6-31G(d,p) basis set. Moreover, potential energy surfaces (PES) were also

performed at B3LYP/MM level to check the reliability of the employed computational technique.

The second contribution of this PhD Thesis, (Contribution II in Chapter 7) corresponds to the study of the same reaction but in the active site of DHAK of *E. coli*, exploring each catalytic step by means of free energy calculations from their corresponding previous PES. *Spline* corrections on the FESs and PES were also conducted at higher level of calculation (DFT). A current debate exist about the most favourable mechanism catalysed by kinases, therefore, in this work the two principal phosphorylation mechanisms discussed in the literature were explored: a *substrate-assisted mechanism*, where no other residue than the ATP substrate participate in the reaction, and the *asp-assisted mechanism*, where the conserved residue Asp109A acts as a catalytic base. A detailed structural and energetic analysis have been done of the located stationary points where the results obtained at low and high level of theory have been compared.

In collaboration with the group of Prof. Eduardo García-Junceda of the Institute of General Organic Chemistry of CSIC (Madrid, Spain), an attempt of tuning DHAK protein to use an inorganic polyphosphate (poly-P) as the phosphoryl donor has been carried out by means of protein engineering techniques in combination with computational methods (Contribution III in Chapter 7). The group of E. García-Junceda synthetized a multi-enzyme with kinase and aldolase activity which requires an ATP regeneration system, sometimes problematic. A very advantageous compound as a phosphoryl donor is the poly-P. Since the DHAK of *C. freundii* does not have activity with poly-P, a directed evolution program was conducted by the group of E. García-Junceda in order to modify the specificity of the DHAK from the ATP to the poly-P. They found mutant forms of the enzyme which presented activity with poly-P, being the DHAK with the single mutation Glu526Lys the one that showed the largest activity. Theoretical calculations based on molecular dynamics simulations (MD) within QM/MM potentials allowed to study the interaction energies in the wild type and the mutant form to analyse the effects caused by the mutation on the binding of the poly-P. Results could be used to explain the activity of the mutant designed by the group of Prof. E. García-Junceda.

Finally, the last contribution of the present PhD Thesis is the exploring of the molecular mechanism of the phosphoryl transfer reaction from poly-P

to Dha in wild type and in the Glu526Lys DHAK mutant from *C. freundii* (Contribution IV in Chapter 7). The final goal of this study was to elucidate whether the Glu526Lys substitution has catalytic effects not only in the binding step but also in the chemical step. PES at PM3/MM level were explored and the location of key stationary points at B3LYP/6-31G(d,p) was carried out considering the *substrate-assisted mechanism* and *the asp-assisted mechanism.* The similar energy barriers obtained in both proteins supported the conclusions derived from our previous study.

Introduction
(Spanish version)

Las proteinas son las moléculas orgánicas más abundantes en las células. Existen en cada una de las partes de los seres vivos puesto que son fundamentales en su estructura y su función celular. Las enzimas son proteínas especializadas en la catálisis de las reacciones biológicas y están entre las biomoléculas conocidas más destacadas debido a su extraordinaria especificidad y poder catalítico.[1] Además, las enzimas son capaces de funcionar en condiciones muy suaves de temperatura y pH. Uno de los requisitos fundamentales que necesitan los organismos para vivir es la capacidad de catalizar reacciones químicas de una manera eficiente y selectiva. Así pues, las enzimas son las responsables de catalizar las reacciones químicas correspondientes a la descomposición de moléculas de nutrientes, la síntesis de macromoléculas biológicas desde simples precursores y la conservación y transformación de la energía química.[2,3] Todo trabajo de investigación sobre las enzimas es de gran relevancia considerando que su presencia y el diferente grado de actividad en los seres vivos puede estar relacionado con enfermedades. La causa, de una enfermedad determinada, dependiendo de las condiciones, puede ser la deficiencia o total ausencia de una o más enzimas (por ejemplo, trastornos genéticos hereditarios). Sin embargo, también podría ser generada por una excesiva actividad enzimática. El diagnóstico puede ser obtenido por medio de medidas de la actividad de las enzimas en el plasma sanguíneo, eritrocitos o muestras de tejidos.[3]

Numerosos trabajos de investigación se han realizado para entender el origen del poder catalítico de las enzimas.[4-9] Una profunda comprensión de los mecanismos enzimáticos, es decir entender la manera en que las enzimas catalizan reacciones químicas, es esencial para analizar los procesos bioquímicos y puede contribuir a producir nuevas medicinas o nuevos catalizadores.[10] Para alcanzar esta comprensión, un aspecto básico a considerar es la comparación de la reacción catalizada con su equivalente no catalizada, es decir, la reacción en disolución acuosa.[10] Cabe señalar que la reactividad química de las moléculas está determinada principalmente por las energías relativas de reactivos, estado/s de transición, posible/s intermedio/s y productos. Estas estructuras críticas se encuentran a lo largo del camino de reacción que conecta reactivos y productos a través de la mínima energía.

La química computacional se ha convertido en una herramienta muy útil para estudiar la reactividad enzimática.[11] Es capaz de proporcionar información estructural y energética sobre compuestos correspondientes a

estados de transición e intermedios y puede hacer perceptible procesos bioquímicos esenciales como el plegamiento de proteínas, el transporte de membrana, la interacción con fármacos y los cambios conformacionales cruciales para el propósito de las proteínas. Estos procesos no pueden observarse directamente a través de las técnicas experimentales.[10,12,14]

Así pues, las simulaciones computacionales de los sistemas enzimáticos pueden contribuir a la comprensión del aumento extraordinario de la constante de velocidad que caracteriza las enzimas así como esclarecer los posibles mecanismos de reacción.[6] Las barreras de energía libre de la reacción obtenidas en las simulaciones pueden compararse con medidas experimentales ayudando a establecer el mecanismo que hay detrás de la cinética experimental. Además, las mejoras en los métodos y arquitecturas computacionales permiten reproducir la reactividad química cada vez mejor.[13] Dado que se pueden emplear diferentes métodos para modelizar las reacciones enzimáticas, la selección del más adecuado es una tarea importante.

En esta Tesis Doctoral se presenta un estudio del proceso catalítico en quinasas, concretamente, en las dihidroxiacetona quinasas (DHAKs). Estas proteínas catalizan la reacción de transferencia de fosfato desde el adenosín trifosfato (ATP) a la dihidroxiacetona (Dha) generando dihidroxiacetona fosfato (Dha-P) y adenosín difosfato (ADP). Las quinasas tienen muchas propiedades y aplicaciones relevantes que hacen interesante su estudio. Se ha elegido una DHAK representativa de cada una de las dos principales clases existentes: la DHAK del organismo *Escherichia coli* (*E. coli*) y la DHAK del organismo *Citrobacter Freundii* (*C. freundii*). La metodología empleada en los estudios presentados en este trabajo está basada en el uso de potenciales híbridos de mecánica cuántica/mecánica molecular (QM/MM).

La primera contribución de la presente Tesis Doctoral es el estudio de la reacción de transferencia de fosfato desde el ATP a la Dha en disolución acuosa (Contribución I en el Capítulo 7). Dicho estudio se llevó a cabo empleando diferentes Hamiltonianos semiempíricos (AM1d, PM3 y PM6) para describir la región QM y el subsistema MM fue tratado por medio de los campos de fuerza clásicos OPLS (para describir una parte del ATP) y TIP3P (para describir las moléculas de agua). Se exploraron superficies de energía libre (FES) a partir de las superficies de energía potencial (PES) previas, para los mecanismos concertado y por pasos, considerando también, el mecanismo asistido por una molécula de agua. Así pues, se

llevó a cabo un estudio estructural, mecanístico y energético de los mecanismos propuestos. Teniendo en cuenta las limitaciones de los métodos semiempíricos para describir este tipo de reacciones, se llevaron a cabo correcciones *spline* sobre las FES con el funcional de la densidad B3LYP y la función de base 6-31G(d,p). Además, se calcularon PES al nivel B3LYP/MM para comprobar la validez de la técnica computacional empleada.

La segunda contribución de esta Tesis Doctoral (Contribución II en el Capítulo 7) corresponde al estudio de la misma reacción pero en el centro activo de la DHAK de *E. coli*, explorando cada paso catalítico por medio de FESs a partir de sus correspondientes PESs. Igualmente, se realizaron correcciones *spline* sobre las FES y PES a nivel (DFT). Dado que existe un debate actual sobre el mecanismo de reacción de las quinasas, en este trabajo se han explorado los dos mecanismos principales discutidos en la literatura: el mecanismo asistido por el sustrato, donde solo el sustrato ATP participa en la reacción y el mecanismo asistido por el Asp, donde el residuo Asp109A actúa como base catalítica. Se ha elaborado un análisis detallado estructural y energético de los puntos estacionarios localizados comparando los resultados obtenidos a los diferentes niveles de cálculo teórico.

En colaboración con el grupo del Prof. Eduardo García-Junceda del Instituto de Química Orgánica General del CSIC (Madrid, España), se ha intentado generar actividad en la proteína DHAK con un polifosfato inorgánico (poly-P) como dador de grupos fosfato por medio de técnicas de ingeniería de proteínas en combinación con métodos computacionales (Contribución III en el Capítulo 7). El grupo del Prof. E. García-Junceda sintetizó una multi-enzima con actividad quinasa y aldolasa que requiere un sistema de regeneración de ATP, a veces problemático. Un compuesto muy ventajoso como dador de grupos fosfato es el poly-P. Como la DHAK de *C. freundii* no presenta actividad con el poly-P, el grupo de E. García-Junceda llevó a cabo un programa de evolución dirigido para modificar la especificidad de la DHAK desde el ATP al poly-P. El grupo encontró varias formas mutantes de la enzima que presentaban actividad con el poly-P siendo la DHAK con la única mutación Glu526Lys, la que presentaba la mayor actividad. Cálculos teóricos basados en simulaciones de dinámica molecular (MD) con potenciales QM/MM permitieron el estudio de las energías de interacción en la enzima nativa y la mutada para analizar los efectos causados por la mutación en la unión del poly-P. Los

resultados pudieron ser empleados para explicar la actividad observada en la enzima mutante diseñada por el grupo del Prof. E. García-Junceda.
Finalmente, la última contribución de la presente Tesis Doctoral es la exploración del mecanismo molecular de la transferencia de fosfato desde el poly-P a la Dha en la DHAK de *C. freundii* nativa y con la mutación Glu526Lys (Contribución IV en el capítulo 7). El objetivo final de este estudio es esclarecer si la sustitución Glu526Lys tiene un efecto catalítico no solo en la etapa de unión del poly-P sino también en la etapa catalítica. Se exploraron PES al nivel PM3/MM y se localizaron los puntos estacionarios al nivel B3LYP/6-31G(d,p) considerando los dos mecanismos, el asistido por el sustrato y el asistido por el aspartato. Los valores similares de energías de activación obtenidos en ambas proteínas confirmaron las conclusiones de nuestro estudio anterior.

2. Objectives

Kinase enzymes catalyse phosphorylation reactions which are implicated in fundamental processes occurring in cells of living organisms. In this PhD Thesis, the phosphoryl transfer mechanisms of DHAKs have been tried to elucidate employing hybrid QM/MM potentials. Possible reaction pathways between Dha and the common phosphoryl donor employed by kinases, ATP, and the cheaper inorganic poly-P have been presented. In order to use the poly-P as a phosphoryl donor, the specificity of DHAK has been modified and the effects caused by a single mutation over the poly-P have been analysed. Since kinases have a conserved active site, the mechanisms here proposed may be suitable for the whole protein family and it can help to understand the catalytic behaviour of these enzymes.
The specific objectives that were proposed and addressed in this PhD Thesis are the following:

- To study different phosphoryl transfer mechanisms from ATP to Dha in aqueous solution with different semiempirical Hamiltonians.

- To determine the most favourable phosphoryl transfer mechanism from ATP to Dha catalysed by DHAK from *E. coli* through the analysis of PESs and their corresponding PMFs employing a computational methodology derived from the study of the reaction in solution.

- In collaboration with protein engineering techniques, to modify the wild-type DHAK from *C. freundii* and to evaluate the catalytic effects of the most active mutant on the binding step of a Poly-P to the protein based on analysis from QM/MM MD simulations.

- To explore the chemical step of the phosphoryl transfer mechanism from poly-P to Dha catalysed by the wild type and the most active mutant of DHAK from *C. freundii* through the evaluation of PESs.

3. Enzymes

3.1. Early discoveries in enzymatic world

Since nineteenth century, scientists have tried to understand the enzymatic behaviour and they have been capable of deepening into the enzymatic world, carrying out the isolation of the enzymes, determining its structure, explaining the role fulfilled and achieving successfully its synthesis.

In 1835, Berzelius published the first general theory about chemical catalysis, where he pointed out that a malt extract known as diastase (nowadays, it is proved that it contains α-amylase) catalysed the hydrolysis of the starch more efficiently than sulphuric acid.[15]

The impossibility of reproducing most of the biochemical reactions in the laboratory led to some authors to suppose that living systems had a "vital power" that allow them to elude the nature laws which govern the inanimate substances. Among these authors was Luis Pasteur who suggested in 1850 that the fermentation process was only possible in alive cells.[3,16] The name "enzyme" (*en*: in, *zyme*: yeast) was first used by Frederik W. Kühne in 1878 with the attempt of emphasizing that there is something in yeast, which is not the yeast itself, capable of catalysing fermentation reactions. In 1897, Buchner demonstrated that molecules promoting fermentation can function even when they are removed from cells; they proved it in the synthesis of ethanol from glucose carried out by a yeast extract free from cells.[3,16] From this point, a huge advance in the biochemistry field was produced; in fact, in 1926, the Nobel Prize J.B. Sumner isolated and crystallize for the first time an enzyme, the urease, showing that the crystals were entirely protein.[17] In the subsequent years, other enzymes were crystallized and founded also to be proteins. During this period J.B.S. Haldane exposed his hypothesis about the enzymatic catalysis, mentioned in section 3.3. In 1960 the first amino acid sequence of an enzyme, the ribonuclease A, was fully described by Stein[18] and in 1967, Phillips D.C. determined the first tridimensional structure of the lysozyme enzyme by X-ray diffraction[19] providing new knowledge to the structural biology.

3.2. General properties of enzymes

An enzyme is usually a protein molecule built up of 20 different species of amino acids linked together to form one or more long chains. Its molecular

weight ranges from ten thousand to hundreds of thousands of Daltons.[20] The long chain(s) must be folded in such a way that a three-dimensional pocket or cleft is created into which the compounds attached, known as substrates, fit in a very precise way.[20] This part of the enzyme where takes place the binding of the substrate and the catalysis is the active site.[21]
Enzymes can be classified into the following six classes, depending on the type of catalysed reaction:[3,16]

1. *Oxidoreductases;* transfer of electrons (hydride ions or transfer atoms).
2. *Transferases;* transfer of specific functional groups.
3. *Hydrolases;* hydrolysis reactions (transfer of functional groups to water).
4. *Lyases;* groups addition to double bonds or formation of double bonds by removal of groups.
5. *Isomerases;* transfer of groups within molecules to generate isomeric forms.
6. *Ligases;* formation of C-C, C-S, C-O or C-N bonds by condensation reactions coupled to adenosine triphosphate (ATP) cleavage.

Each enzyme has a systematic name and a four-part classification number. [3,16] The official systematic name is formed adding the suffix –ase to the name of its substrate(s) or to a word or phrase, preceded by the name of the substrate, specifying the type of reaction catalysed. An example would be the enzyme Dihydroxyacetone Kinase. The classification number from the Enzyme Commission for Dihydroxyacetone Kinase is EC 2.7.1.29, where:

the first number, 2, indicates the class name: transferase;
the second number, 7, means the subclass phosphotransferase which involves the transferring phosphorous-containing groups;
the third number, 1, denotes that it is a phosphotransferase with an alcohol group as acceptor;
the fourth number, 29, indicates that it has a glycerone as the phosphoryl group acceptor;

There are enzymes that require cofactors to be active, which are non-proteic components either one or more inorganic ions, such as Mg^{2+}, Zn^{2+}, Mn^{2+}, Cu^{2+}, etc., or organic or metallo-organic molecules called coenzymes. Some enzymes require both.[3,16] A coenzyme can also act as a co-substrate when it is only transiently associated with the enzyme (e.g., NAD^{+}, $NADP^{+}$). A

coenzyme or metal ion can act as a prosthetic group when it is permanently bounded to the enzyme, often by covalent bonds (e.g., an heme group). A complete catalytically active enzyme, bounded with its cofactor, is called a holoenzyme whereas, the inactive protein without its cofactor, is denominated apoenzyme. Enzymes undergo chemical changes due to its participation in reactions and they must return to its original state in order to complete the catalytic cycle.
The most remarkable properties of enzymes as catalysts are:[16]

- *Catalytic power:* The rate of enzymatic reactions is normally from 10^6 to 10^{12} higher than the corresponding non-enzymatic reactions.

- *Specificity:* Enzymes usually have a much higher degree of specificity with respect to their substrates (reactants) and products than their chemical catalysts, which also means that enzymatic reactions rarely have side products.

- *Milder reaction conditions:* Enzymes commonly work under neutral pH, mild temperature and atmospheric pressure whereas efficient chemical catalysis in the laboratory often requires extreme conditions.

- *Capacity for regulation:* The activity of some enzymes can vary in response to concentration of substances different to their substrates. These regulatory mechanisms include variation of the quantity of synthetized enzymes, allosteric control and covalent modification of the enzymes.

Enzymes catalyse reactions at rates that are often incomprehensibly faster than non-enzymatic counterparts. It is important to highlight that the purpose in mechanistic enzymology is to determine the exact reaction mechanisms and the physicochemical phenomena contributing to this enormous rates reached by enzymes.[22]
It is also interesting to stand out that both pH and temperature have an effect on enzymatic catalysis; most of the enzymes have a characteristic pH, which is usually neutral, where they have their maximum activity. The relationship between pH and activity depends on the acid-base behaviour of the enzyme and substrate.[3,23] Due to the fact that at high temperatures enzyme may become denatured, it exists an “optimal” temperature which is an equilibrium between both factors.

Nevertheless, there are a sort of enzymes which are able to evolve at higher temperatures. They are classified as thermophiles, characterized for growing up at a range of temperatures between 50 to 70ºC and hyperthermophiles, present at temperatures higher than 80ºC. These proteins usually are in places such as thermal waters and they exist even at temperatures up to 130ºC.[16,24] Curiously, these enzymes have many common metabolic pathways with the called mesophilic enzymes, which are those with an optimal growth between temperatures from 24 to 40 ºC. One of the differentiating key, responsible of the stabilized protein structure of thermophilic proteins is the presence of an overabundance of saline bridges settled forming nets. Other structural features causing of this thermostability are; van der Waals interactions, hydrogen bonds, disulphide bonding, the incremented size of the hydrophobic nucleus of the protein or in the interface between its subunits or domains and the improvement of the quality of packing.[24,25] Study of these properties is important to understand correctly the evolutionary process of enzymes and the energetics that control the protein folding, recognition and stability.[24]

3.3. Enzymatic catalysis

The fundamental role of enzymes is the enhancement of reaction rates. To understand the way in which this enhancement is produced is necessary a kinetic description of their activity; that is, the determination of the rate and its evolution in response to alterations of experimental parameters.
In 1913, Leonor Michaelis and Maud Menten[26] expanded the idea that the formation of the enzyme-substrate (ES) complex as a necessary step in enzymatic catalysis, into a general theory of enzyme action that follows the catalytic scheme (for a single-substrate enzyme catalysed reaction):[3,21]

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P \qquad (3.1)$$

In this mechanism the enzyme (E) is combined reversibly with the substrate (S) resulting in the ES complex in a relatively fast step. This complex reacts in order to form a product (P) and the releasing of the enzyme in a slower step. k_1 and k_{-1} are the rate constants for the direct and reverse processes of the formation of the ES complex. k_{cat} is the rate constant for the step of products

formation that is supposed to be irreversible. The reaction rate for a one substrate reaction catalysed enzymatically is:

$$v_0 = \frac{v_{max}[S]}{K_M + [S]} \tag{3.2}$$

Equation (3.2) is known as the Michaelis-Menten equation.

Initially, it was thought that exclusively the formation of the ES complex was the reason of this lowering. In fact, in 1894 Emil Fischer,[27] basing on this idea, proposed that enzymes were structurally complementary to their substrates, both rigid, and fitted like a lock and a key (see Figure 3.1a). A more recent hypothesis was proposed by Michael Polanyi[28] (1921) and Haldane (1930) and is based on the view that an enzyme is flexible, capable of changing. They pointed out that the active site of the enzyme adopts an optimal conformation to interact with the substrate, but only when the substrate is bounded (see Figure 3.1b). The modern concept of enzymatic catalysis follows this hypothesis and was elaborated by Linus Pauling in 1946[29] who stated that an enzyme should be complementary to the transition stated for the catalysis of reactions, that is, optimal interactions between enzyme and substrate take place exclusively in the transition state. An enzyme with a pocket complementary to the reaction transition state helps to make unstable the substrate, promoting the catalysis reaction. The increase in the free energy required to modify the substrate is compensated by the binding energy released from the weak interactions between the enzyme and substrate in the transition state.[3]

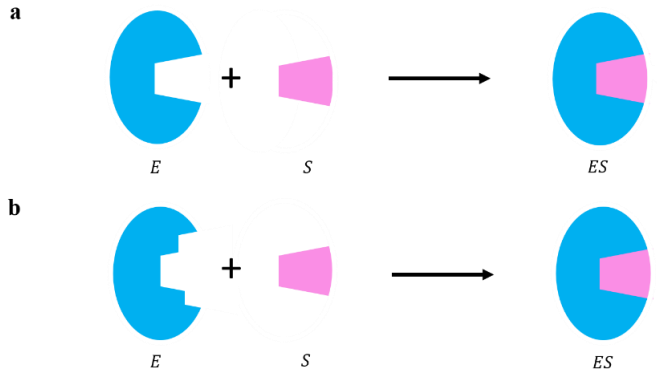


Figure 3.1. Schematic representation of lock-and-key model (a) and induced-fit model (b).

In Figure 3.2 is displayed a free energy profile of a catalysed reaction and a non-catalysed one.[30] The former corresponds to the kinetic scheme proposed by Michaelis and Menten presented in equation (3.1).

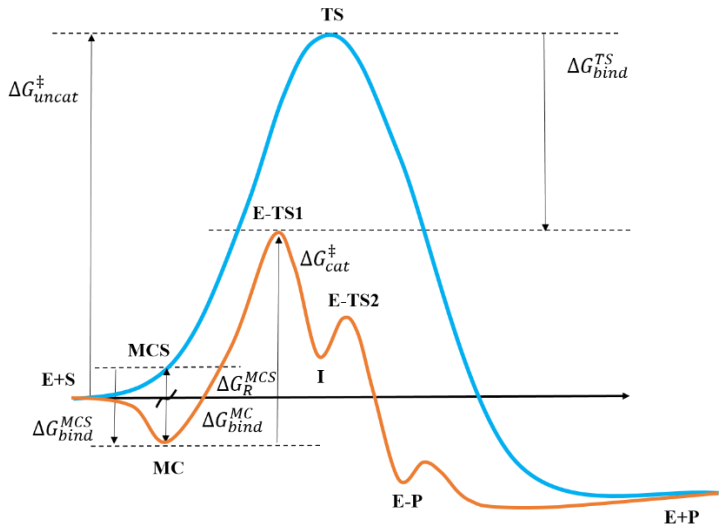


Figure 3.2. Free energy diagram of two-step reaction catalysed by an enzyme (orange line) compared to the non-catalysed reaction (blue line). Figure adapted from reference[30].

The non-catalysed reaction occurs in just one step, with an activation free energy of $\Delta G^{\ddagger}_{uncat}$. In contrast, the reaction catalysed by the enzyme consists in two steps: first, the enzyme substrate, known in this example as the Michaelis-Menten complex (MC), is formed losing binding energy, ΔG^{MC}_{bind}. Secondly, the chemical reaction takes place with an energetic cost of $\Delta G^{\ddagger}_{cat}$. This value is smaller than $\Delta G^{\ddagger}_{uncat}$ and it can be related to the reaction rate, k_{cat} by the Transition State Theory (this theory is detailed in section 5.5.):

$$k_{cat}(T) = \frac{k_B T}{h} exp\left\{\frac{\Delta G^{\ddagger}_{cat}(T)}{RT}\right\} \qquad (3.3)$$

The catalytic power of an enzyme is usually expressed by the ratio k_{cat}/k_{uncat}, that is, it is proportional to the difference between the activation free energy of the uncatalysed reaction and the catalysed one. According to the

Figure 3.2 the reduction of the activation free energy can be expressed in terms of the binding energy of reactants and TS:

$$\Delta G^{\ddagger}_{uncat} - \Delta G^{\ddagger}_{cat} = \Delta G^{MC}_{bind} - \Delta G^{TS}_{bind} \quad (3.4)$$

If the enzyme speeds up the reaction, the left term of equation (3.4) should be positive, so ΔG^{TS}_{bind} should be greater in its absolute value than ΔG^{MC}_{bind} (it must be bear in mind that binding energies are always negative in value). This means that the enzyme presents a larger affinity for the transition state than for the reactant state, supporting the theories that assume that the catalytic role of the enzymes is based on *the stabilization of the TS* versus the MC, firstly stated by Pauling[29] as indicated above. Warshel and co-workers[31-33] pointed out that this stabilization is due to an electric field arranged by the active site to adjust the charge distribution in the TS.

However, it exists a thinking that gives priority to the formation of the Michaelis complex in the catalysis. This is better understood if it is defined an imaginary and similar structure to the MC but in solution, MCS. The free energy needed to arrive from reactants in solution to the MCS structure, ΔG^{MCS}_{R}, and the binding free energy of the MCS complex, ΔG^{MCS}_{bind}, would follow this equation:

$$\Delta G^{MC}_{bind} = \Delta G^{MCS}_{R} + \Delta G^{MCS}_{bind} \quad (3.5)$$

If we replace equation (3.5) in equation (3.4):

$$\Delta G^{\ddagger}_{uncat} - \Delta G^{\ddagger}_{cat} = \Delta G^{MCS}_{R} + \Delta G^{MCS}_{bind} - \Delta G^{TS}_{bind} \quad (3.6)$$

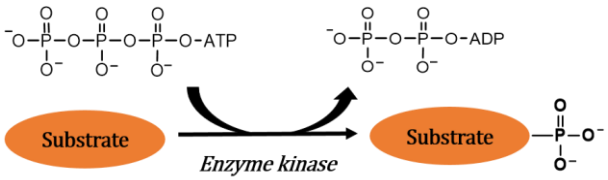
Since ΔG^{MCS}_{R} is always positive, a high catalytic power ($\Delta G^{\ddagger}_{uncat} - \Delta G^{\ddagger}_{cat} > 0$) can be obtained even when the enzyme shows greater affinity to the MCS structure than to TS. Then, in this case it is supposed that the enzyme speeds up the reaction by means of *preorganization of the substrate*. Theories that follow this idea are based on the fact that it is produced an important change in the substrate geometry since there is a transition from separated and solvated species towards a spatial reorganization where they are much closer. This means that ΔG^{MC}_{bind} may contribute essentially in the catalysis.

Apart from these theories, it should be considered other effects which, especially during the last years, has mentioned in the literature to contribute in

a relevant way in enzymatic catalysis such as, dynamical effects or tunnelling.[6] These effects are discussed in detail in section 5.5.1.

4. Kinases:
Dihydroxyacetone kinases

A kinase is an enzyme that catalyses the transfer of the terminal phosphoryl group from ATP (or guanosine triphosphate, GTP) to a substrate containing an alcohol, amino, carboxyl, or phosphate group as the phosphoryl acceptor, in a process called phosphorylation (see Scheme 4.1).[34,35] These enzymes normally hold two Mg^{2+} ions placed into the active site which assist the phosphorylation catalysis (the role of Mg^{2+} ions will be discussed in the section 4.2).[36,37]



Scheme 4.1. Schematic representation of the phosphorylation process by a protein kinase, concomitant with the release of adenosine diphosphate (ADP).

The first protein kinase to be purified was phosphorylase kinase in 1959 by Krebs et al[38] followed in 1968 by a protein-serine kinase, the cAMP-dependent protein kinase[39]. In 1979 the modification of proteins by phosphorylation on tyrosine residues was discovered through the study of tumour viruses.[40] The following year, Hunter and Sefton found that the transforming protein of Rous sarcoma tumour virus, pp60vsrc, has tyrosine phosphorylation activity and thus, was classified as a kinase. This was an important discovery since it implied that deregulated protein tyrosine phosphorylation may be important in tumorigenesis.[41] In 1980 the number of identified protein kinases starting to increase drastically being an important factor the recognition that all the protein kinases have a similar sequence in their catalytic domains.[42] The first evidence of this similarity came when the catalytic subunit of cAMP-protein kinase was related with pp60vsrc over a region of approximately 300 amino acids which was an unexpected revelation.[2,43] The first kinase enzyme structure crystallographically solved and analysed was the cAMP-dependent protein kinase (PKA), in 1991,[44,45,39] which provided a framework for the entire family and sometimes has been used as a model for the rest of the family members.[46-48]

Phosphorylation reactions are involved in many processes taking place in the cells of living organisms, contributing in an essential way in propagation and signal transduction.[49-52] These reactions usually result in a functional

modification of the target protein or substrate by changing their activity, cellular location, or association with other proteins. Phosphorylation processes are capable of producing contrary effects such as the increasing or decreasing of enzyme activity and stabilization or destabilization of a protein. In addition, some phosphorylation sites of a protein can be stimulatory while others are inhibitory.[37] An evidence of the relevance of kinases is the fact that there are more than 500 protein kinases encoded in the human genome. In addition, it has been found of about 540 kinases in mice and 122 in yeast, representing one of the largest protein family.[53,54]

Numerous kinases have been detected to be closely involved in cancer progression, inflammation and autoimmune disorders[55-57] and the target distribution within the human "druggable genome" shows that these enzymes represent an attractive field of activity for medicinal chemistry (see Figure 4.1).[54,58-61] In fact, it has been demonstrated that about 75% of the kinome are in principle druggable so they could be ideally used with a therapeutic benefit to patients.[57] The proof is that some small-molecule kinase inhibitor drugs have been approved already for the treatment of cancers.[57,62,63]

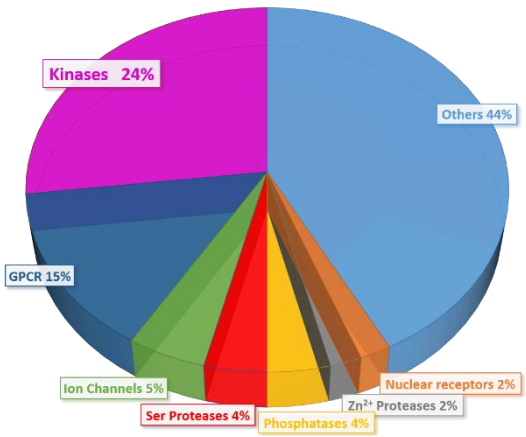


Figure 4.1. Schematic presentation of the target family distribution within the human "druggable genome". Figure adapted from reference [49].

Members of protein kinase family has a significant common feature which is the sharing of a highly conserved catalytic center, as indicated above.[49,52,64-66] This characteristic makes possible that one inhibitor molecule can bind to several kinases, preventing from the usual signaling needed for normal cellular behavior. However, it may be also feasible that the capability of interfering in

a multitude of kinases is beneficial in cancer treatments, provided that the inhibitor is specific to cancerous cells.[49,54] The sharing of this conserved catalytic domain is also responsible of a common catalytic mechanism.[54]
Mainly, eukaryotic protein kinases can be divided generally into ten groups basing on the sequence identity:[57,67]

1- protein kinase A, G, and C family altogether called as AGC kinases.

2- CMGC group of kinases formed by: Cyclin-dependent kinases (CDKs), mitogen-activated protein kinases (MAP kinases), glycogen synthase kinases (GSK) and CDK-like kinases.

3- Calmodulin/ Calcium regulated kinases (CAMK).

4- Casein Kinase 1 (CK1) group.

5- STE group of kinases including Mitogen-activated protein kinases (MAPK) cascade families, which are specific to serine, threonine and tyrosine aminoacids.[68]

6- Tyrosine kinases (TK).

7- Tyrosine kinase like (TKL) generally phosphorylating serine/threonine residue.

8- Receptor Guanylate Cyclases (RGC) group of kinases.

9- Protein kinase like (PKL) proteins having fold similar to kinases.

10- Atypical protein kinases (aPK); consist of proteins which do not share clear sequence similarity with other kinases but are functioning kinase like.

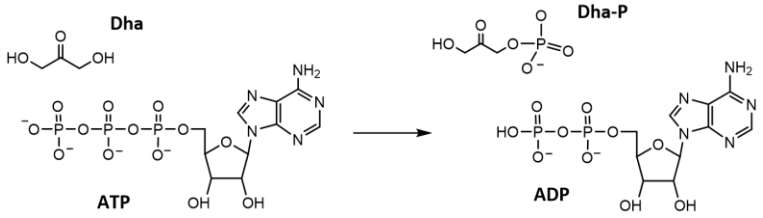
Cheek and co-workers provided a more extensive classification where kinases are classified in 12 groups, by family and fold group.[34]

1- Protein S/T-Y kinase/atypical protein kinase/ lipid kinase/ ATP-grasp.
2- Rossmann-like.
3- Ferredoxin-like fold kinases.
4- Ribonuclease H-like.
5- Triose phosphate isomerase (TIM) β/α-barrel kinase.
6- Galactokinase, Homoserine kinase, Mevalonate kinase and Phosphomevalonate kinase which are the GHMP kinases.
7- Aminoimidizole ribonucleotide (AIR)-synthetase like.

8- Riboflavin kinase.
9- Dihydroxyacetone kinase (DHAK).
10- Putative glycerate kinase.
11- Polyphosphate kinase.
12- Integral membrane kinases.

In this classification is included the DHAK. This enzyme represents a new unique kinase fold recently characterized and the exploration of its reaction mechanism has been the main subject of this PhD Thesis.

DHAKs phosphorylate dihydroxyacetone (Dha) converting it in Dha phosphate (Dha-P), as shown in Scheme 4.2:[69,70]



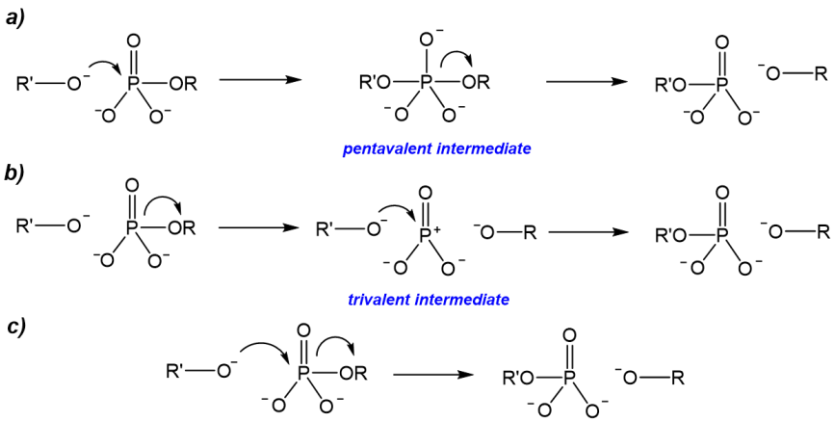
Scheme 4.2. Schematic representation of the phosphoryl transfer from ATP to Dha, generating ADP and Dha-P.

Dha-P is an intermediate for the synthesis of pyruvate[70,71] and it is a very important specie for C-C bond formation since it is used as substrate donor in several enzyme-catalysed aldol reactions.[72,73] DHAKs are divided into two classes, those using the phosphoenol pyruvate carbohydrate (PEP): phosphotransferase system (PTS) to provide the phosphoryl group, in most bacteria, or those using ATP as the phosphoryl donor in animals, plants and some bacteria.[69,70,74-78]. In this PhD Thesis it has worked with the main DHAKs of each class. Nevertheless, before deepening into them, the general possible phosphorylation mechanisms and the role of Mg^{2+} ions in kinase enzymes will be presented and discussed.

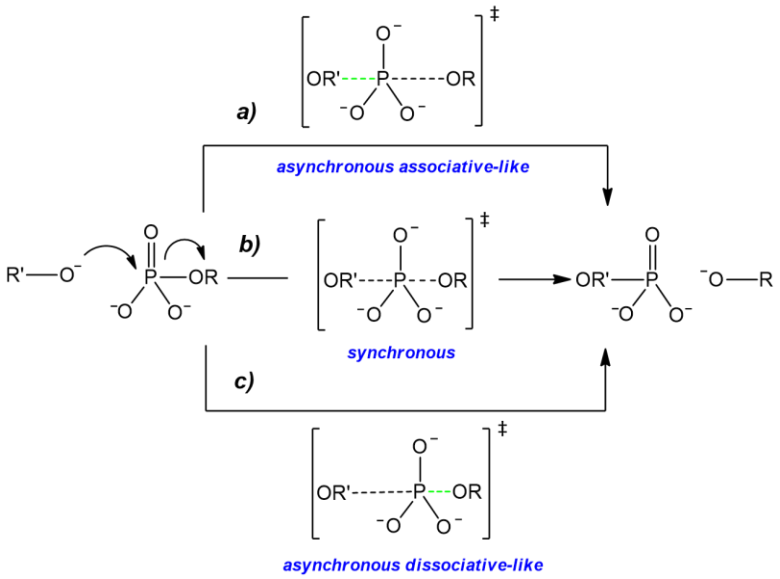
4.1. Mechanisms of phosphate transfer in protein kinases

The accessibility of low-lying d-orbitals in the phosphorus atoms makes possible their use in bonding interactions. In consequence, the phosphorus atom is easily capable of expanding its coordination number in order to give pentavalent compounds.[79-83] Bearing in mind this, phosphorylation reactions

can consider to occur through two likely mechanisms:[81,84-86] associative, where nucleophilic attack is produced before the departure of the living group, and dissociative, where the living group departure occurs previously to the nucleophilic attack. These reactions may take place within several steps in a process called stepwise pathway or by means of only one step that is the case of a concerted pathway. A stepwise pathway implies the presence of more than one transition state structure and then, the existence of intermediates which are pentavalent structures in associative processes (see Scheme 4.3a)) or trivalent compounds in dissociative procedures (see Scheme 4.3b)). A concerted pathway has one transition state, which indicates the absence of intermediate species (see Scheme 4.3c)). Nevertheless, the description of a concerted pathway is more complex. A concerted transition state can be synchronous; with similar amount of bounding for the nucleophile and the living group (see Scheme 4.4b)) or asynchronous. Simultaneously, an asynchronous transition state can be associative-like mechanism (see Scheme 4.4a)), with a larger degree of bond formation to the nucleophile than cleavage of the bond to the leaving group or dissociative-like mechanism with more bond cleavage than bond formation (see Scheme 4.4c)).

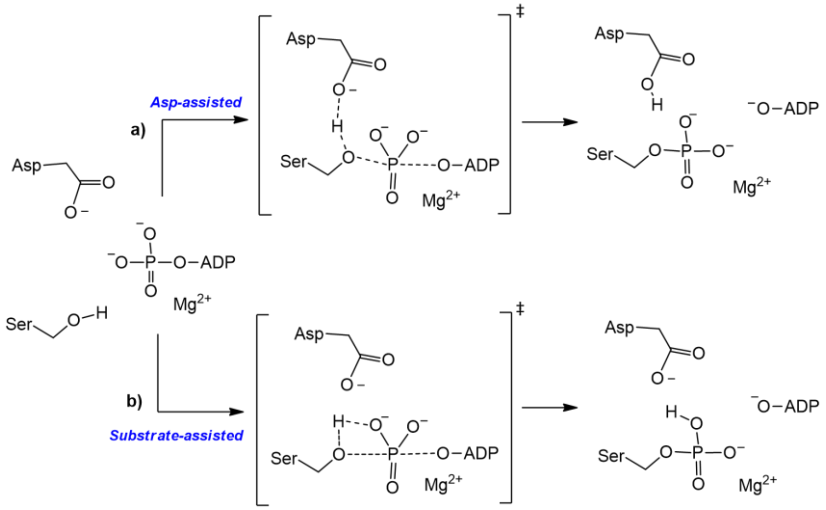


Scheme 4.3. Schematic representation of the general mechanisms in phosphate transfer reactions. a) Stepwise associative mechanism with the formation of a pentavalent intermediate. b) Stepwise dissociative mechanism with the formation of a trivalent intermediate. c) Concerted mechanism. Figure adapted from reference[84].



Scheme 4.4. Schematic representation of the different possibilities in a concerted phosphate transfer mechanism. a) Concerted asynchronous transition state proceeding in an associative-like manner. In green has been highlighted the bond formation with a larger degree of union than the cleavage bond. b) Concerted synchronous transition state. c) Concerted asynchronous transition state proceeding in a dissociative-like manner. In green has been highlighted the cleavage bond with a larger degree of union than the formation bond.

It is very common to find research works relatives to the phosphoryl transfer from ATP, the phosphoryl donor molecule in protein kinases, to a serine or tyrosine residues. Authors usually distinguish between phosphorylation reactions where the conserved Asp residue acts as a base activating the serine, or tyrosine residues for the nucleophilic attack or where this activation is caused by the ATP substrate. The former process is categorized as an *asp-assisted mechanism* and the latter is called a *substrate-assisted mechanism*. Both processes are displayed in Scheme 4.5.



Scheme 4.5. Schematic representation of the two phosphoryl transfer mechanisms proposed in protein kinases. In the scheme, the γ-phosphate of the ATP substrate is transferred to a Ser residue. a) *Asp-assisted* mechanism where the activation of the Ser residue for the nucleophilic attack is produced by a conserved Asp. b) *Substrate-assisted* mechanism where the activation of the Ser residue for the nucleophilic attack is produced by the ATP substrate. Both mechanisms are displayed in a concerted process for clarity. Figure adapted from reference[52].

There is an extensive debate which still remains open about the most favorable phosphoryl transfer mechanism in kinase enzymes. Some authors consider the *asp-assisted mechanism* as the most favorable one[48,52,87-90] while others support the *substrate-assisted mechanism* as the most likely.[65,91,92]

Asp-assisted mechanism

Regarding to the studies in favor of the *asp-assisted mechanism*, it will be presented research works in the cAMP-dependent protein kinase (or PKA), the cyclin-dependent kinase (CDK2) and the insulin receptor kinase (IRK).
There are several computational studies in cAMP-dependent protein kinase (or PKA), an enzyme that catalyse the γ-phosphoryl transfer from ATP to a Ser or Thr residue of a substrate peptide.[48,87-89] For example, Diaz and Field[87] conducted classical MD simulations of the catalytic subunit of PKA, Mg_2ATP and a 20-residue peptide substrate, using the CHARMM force field. They showed that the interaction between the carboxylate group of Asp residue and

the nucleophilic Ser was dynamically stable. Quantum mechanical (QM) calculations were also carried out. In order to calibrate the methodology used in the QM study, proton affinities for acetate and methyl phosphate dianion interacting with the Mg^{2+} ions were measured at semiempirical, *ab initio* and Density Functional Theory (DFT)[93] levels (properties of these QM methods will be described in section 5.1.1). Subsequently, they performed QM B3LYP[94,95]/6-31G(d)[96] calculations on a cluster model of a system formed by 92 atoms containing a methyltriphosphate, two Mg^{2+} ions, three water molecules coordinated to the ions and the side chain of Asp166, Lys168, Asn171 and Asp184 residues. Authors located minima and transition state structures (TSs) from a B3LYP/6-31G(d) potential energy surface (PES) which were refined by means of single point calculations at B3LYP/6-31G(d,p) level. They also evaluated the effect of the environment polarity (protein-solvent) performing single-point self-consistent reaction field (SCRF) on the gas-phase QM structures of the stationary points and made electrostatic calculations. Their results revealed as an energetically more favorable the TS where the Asp is ready to accept the proton resulting thus, in an *asp-assisted mechanism*.

A similar catalytic mechanism was suggested later, by Cheng and co-workers.[88] They performed MD and Quantum Mechanical/Molecular Mechanical (QM/MM) calculations. These QM/MM calculations were done in two PKA-ligand complexes. In particular, they employed the pseudobond ab initio QM/MM approach, applying two partition schemes where the difference was whether the ATP was fully included. Therefore, 75 atoms were involved in the QM region for the bigger system while the smaller one enclosed 49 atoms, including the triphosphate part of the ATP. The rest of the residues of the QM region were the same in both systems; side chains of Ser substrate, Asp166, Lys168 and the two Mg^{2+} ions. For searching the stationary points, authors employed the reaction coordinate driving method. Single point calculations at Moller-Plesset (MP2)[97] and B3LYP levels of theory with the 6-31+G(d)[98] basis set were done of these stationary points, from the previous optimized structures at B3LYP/6-31G(d). They also performed the same calculations excluding the Lys168 from the QM part. Their theoretical results suggest the role of Asp166 as a catalytic base, describing the *asp-assisted mechanism* as principally dissociative and the required contribution of Lys168 to stabilize the TS. In addition, they conducted MD using the AMBER force field, of the wild type and mutated forms where the conserved Asp and Lys residues were changed. They founded by means of the analysis of some distances that effectively, the role of Asp was acting as a catalytic base instead

of as a structural anchor to maintain the configuration of the active site. In fact, the latter was suggested to be the role for the Lys residue, keeping the ATP and the substrate peptide in a proper conformation for the nucleophilic attack.

Another examples where the *asp-assisted mechanism* has been analyzed in PKA are the more recently theoretical works of Pérez-Gallegos et al.[48,89] They explored potential energy profiles of the phosphoryl transfer mechanism using QM/MM hybrid methods with DFT (B3LYP/6-31+G(d)) and MP2 functionals. Authors studied the phosphoryl transfer in two different substrates; the heptapeptide synthetic substrate, Kemptide[48] and the 20-residue, SP20[89] substrate. This comparative analysis allowed them to show that the mechanism is not substrate dependent, finding the dissociative mechanism more favorable than the associative one. It should be pointed out that the authors classified the associative mechanism as a one-step reaction where the transition state corresponds to the γ-phosphoryl transfer from ATP to the Ser residue of the peptide substrate where the nucleophilicity of this Ser is increased by the previously proton transfer from its –OH group to one of the oxygen atoms of the phosphate group of the ATP. The dissociative mechanism were defined as a two steps reaction; firstly, the phosphoryl transfer to Ser residue of the peptide substrate takes place where, in this case, the nucleophilicity of the residue is enhanced by the conserved Asp. Consecutively to the phosphoryl transfer, occurs the proton migration to Asp but without potential energy barrier, although an intermediate was located. Secondly, the Asp residue give back the proton to phosphorylated product. In the two research works presented, the definition of the QM region as well as the computational methods employed by the authors were the same. The QM region contained 65 atoms which involved the two Mg^{2+} ions, the phosphate groups of the ATP molecule, three crystallographic water molecules and the side chain of residues Asn171, Asp184, Lys168, Asp166 and $Ser17_{Kemptide}$ substrate in the case of the first study or $Ser21_{SP20}$ substrate in the second work. The MM region was treated by means of the CHARMM force field. Moreover, single point energy calculations were also performed at the MP2/cc-pVTZ[99] level to describe the QM region while the MM subsystem was treated within the CHARMM force field.

There are also different works in other kinase enzymes where the *asp-assisted mechanism* was also supported.[52,90] For example, in the cyclin-dependent kinase (CDK2), a protein that catalyses, as well, the phosphoryl transfer from ATP to a Ser or Thr residue. Smith et al.[52] analysed the reaction by means of free energy calculations employing the QM/MM methodology. Specifically,

they used the pseudobond ab initio QM/MM approach and they performed Potentials of Mean Force (PMFs) for obtaining the free energies and perturbation calculations in order to estimate the influence of surrounding residues on the reactant and TS structures. In their computational model, the QM region contained 48 atoms including the Mg^{2+} ion (the active site of CDK2 hold one single ion), the triphosphate part of ATP, the side chain of residues Asp127, Lys129 and the substrate Ser. This QM subsystem was treated at the B3LYP/6-31G(d) level of theory while the MM region was modeled using the AMBER simulation package. According to their theoretical results, the conserved Asp residue acts as an acid-base in a concerted dissociative-like process within a single metaphosphate-like transition state.

Another kinase where the *asp-assisted mechanism* was observed is the IRK. To understand the role of the autophosphorylations of this enzyme resulting in its activated form, Ojeda-May et al[90] evaluated free energies with QM/MM methods and performed MD simulations. Their simulations displayed a catalytic *asp-assisted mechanism* taking place in a stepwise manner where the conserved Asp abstracts the proton from the Tyr residue followed by the phosphate transfer. The calculations were done in systems of the active conformation and the inactive one of the kinase enzyme. MD simulations were carried out using the TIP3P potential for water molecules and the CHARMM force field for describing the rest of the system. Free energy calculations using the umbrella sampling method and the string method and 2-D PMFs were done. Authors employed two types of QM/MM potential energy functions for these calculations. The first one was the multiscale *ab initio* QM/MM method where the B3LYP functional was used as a high-level QM theory and the semiempirical AM1 as a low-level Hamiltonian. For the B3LYP calculations the basis set used were the following; 6-31G(d) for P atoms, 6-31G for H and O atoms and 3-21G for C, N and Mg atoms. The second type was the semiempirical AM1/d-PhoT (AM1d)[100] QM/MM potential. The QM subsystem established included the triphosphate part of ATP, the side chains of the substrate Tyr and the residues Asp1132, Asn1137 and Asp1150, the two Mg^{2+} ions and six water molecules, comprising a total of 76 atoms. In the case of DFT calculations, some atoms were removed from the QM part, thus, in this case, the QM region contained 56 atoms. Authors found that the increase of the catalytic rate in this enzyme is due to the lowering of the free energy barrier of the proton transfer step. MD simulations supported this theory showing that IRK phosphorylations affected the protein dynamics of the enzyme before the proton transfer to Asp with a smaller effect after the proton is transferred.

Substrate-assisted mechanism

As mentioned previously, there are likewise research works distinguishing the *substrate-assisted mechanism* as the most favorable catalytic pathway.[65,91,92,101] Here it will be presented theoretical studies in PKA, CDK2 and galactokinase (GALK) related with the finding of this mechanism.

Some studies in the kinase PKA supported this mechanism in contrast to those mentioned above.[65,91] An example is the theoretical work carried out by Hart et al.[65] They affirmed that there are experimental data where is questioned whether Asp is basic enough to take the proton from the Ser residue. By means of performing PESs using the QM/MM methodology, they obtained higher potential energy barriers when the proton was transferred from Ser to Asp. The QM part of the computational model was formed by 46 atoms involving the side chains of the Ser substrate and the residues Asp152 and Lys154, the triphosphate part of the ATP and the two Mg^{2+} ions. This subsystem was described using the PM3 level of theory while the MM part was treated by means of the AMBER force field. Authors obtained a concerted transition state where the transfer of the proton to the oxygen atom of ATP and the phosphoryl transfer to the Ser residue occurred in a concomitant way. The Asp residue was suggested to have an alternative role consisting in stabilize the product of the reaction.

Hutter et al.[91] supported as well the *substrate-assisted mechanism* in the same enzyme by means of semiempirical AM1 molecular orbital computations. The quantum mechanical model of the active site was formed by the entire molecule of ATP, the Ser substrate, the two Mg^{2+} ions with their coordination shells and six residues considered to be crucial for the phosphoryl transfer; Gly55, Lys72, Asp184, Asn171, Lys168 and Asp 166 resulting in a total of 123 atoms. Authors performed PESs of the wild-type and a mutant model where the conserved Asp166 was replaced by an Ala finding similar energy profiles. Therefore, their results showed that Asp has not an important role in the phosphoryl transfer itself, however it is important for keeping the configuration of the active site. Moreover, the charge group of Lys168 was changed by a neutral one detecting that this residue was needed to electrostatically stabilize the reaction.

The *substrate-assisted mechanism* was also founded in theoretical studies related with the kinase CDK2[92]. De Vivo and co-workers[92] reported analysis at DFT level in gas phase, classical MD and Car-Parrinello QM/MM simulations concluding that the Asp residue located at the active site plays a

structural role while the abstraction of the proton from the Ser residue is due to the ATP molecule. In the study authors considered four QM models and each model contained the triphosphate moiety, the Mg^{2+} ion, nine crystallographic water molecules and the side chains of the following residues where each one is completed with a CH_3 group: Asp127, Asp145, Asn132, Lys33, Lys129 and Thr165. The models differed in four different protonation states which are; Asp127 deprotonated, Asp127 protonated, triphosphate moiety protonated at one oxygen atom of β-phosphate and triphosphate moiety protonated at one oxygen atom of γ-phosphate. Geometry optimizations of these models were made with the B3LYP functional. For the Car-Parrinelo QM/MM calculations the QM region defined included the triphosphate moiety of ATP, the Mg^{2+} ion and the side chain of Lys129 and Ser substrate involving a total of 32 atoms. The QM subsystem was treated with the BLYP functional while the MM part was described through the AMBER force field. To follow the phosphoryl transfer reaction profile, the so-called blue-moon ensemble simulations were performed in the first model proposed (Asp deprotonated) constraining at different values of the distinguished reaction coordinate chosen. Authors calculated the activation free energies using thermodynamic integration. Results allowed them to support a *substrate-assisted mechanism* through a concerted associative-like TS. It should be remarked that Smith et al. related this conclusion to the exclusion of the Asp from the QM region in the QM/MM simulations.

Huang et al.[101] located the *substrate-assisted mechanism* in GALK where the phosphoryl group is transferred from the ATP to the oxygen of one OH group of the substrate galactose. QM/MM simulations within a two-layer ONIOM method were carried out. The QM region involved ATP molecule, galactose substrate, the Mg^{2+} ion, two water molecules and the truncated residues Asp186, Arg37, Ser142, Glu174 and Arg228 resulting in a total of 49 atoms. This QM region was treated with B3LYP-D/6-31G(d) while the MM region was described by means of the AMBER force field. A PES was explored by the adiabatic mapping method to study the reaction pathway, obtaining a dissociative phosphoryl transfer mechanism by a trigonal-bypiramidal TS. MD simulations of the system inside a box of water molecules were also conducted. The TIP3P potential was used to describe the water molecules while the AMBER force field was employed for treating the rest of the system. In addition, MD simulations of five mutants of GALK were likewise performed in order to analyse the effects of these mutations in the conformation of the active site. Results showed that there were some residues responsible of

restricting the mobility of ATP molecule while others residues were found to be important to stabilize the negative charges during the phosphoryl transfer process .

4.2. Role of Mg^{2+} ions in phosphoryl transfer reactions

The indispensability in many enzymatic systems of the Mg^{2+} ions has encouraged several authors to determine its specific roles in catalysis. Normally, kinases require two magnesium ions to catalyse the phosphoryl transfer, however it has been discovered that the presence of the second metal ion may provoke different effects. Thus, some kinases are activated with a unique Mg^{2+} and the additional one might be required for complete activation, can produce inhibition or may participate in structural stabilization.[36,102,103] In the literature, it has been discussed about the different functions that these ions can accomplish in the active site of kinases and, in phosphoryl transfer reactions in general. They usually contribute in the ATP binding and in the acceleration of the chemical reaction.[36,102] In particular, experimental studies of Bastidas and co-workers in PKA[102] allow them to characterize the role of Mg^{2+} ions. In the experimental procedure followed by the authors, the catalytic subunit of the enzyme was crystallized with the 20 amino acid substrate SP20 and with the nonhydrolyzable analog of ATP, adenosine-5'-(β,γ-imido) triphosphate (AMP-PNP). Two crystal structures were obtained corresponding to two products proving that the presence of both metal ions in the active site prevent from the release of ADP after the phosphate transfer reaction, due to the affinity with the molecule. Then, only one Mg^{2+} cation would remain in the active site while the other one is expelled after phosphate transfer is produced. Even so, the authors have suggested that the presence of both metal ions was essential for an efficient phosphate transfer catalysis. The same conclusions were suggested through the experimental and theoretical studies of Jacobsen et al. in CDK2[36] by means of enzyme kinetics, crystallography and MD simulations. The steady state kinetics of CDK2 was analysed varying concentrations of $ATP{\cdot}Mg^{2+}$ with fixed concentrations of Mg^{2+}. The kinetic data proved that the kinase required two catalytic ions being the second ion greatly cooperative with the binding of ATP and with the other ion. In addition, the interaction between ADP and the phosphorylated kinase were analysed from two crystal structures; one with two Mg^{2+} ions coordinating the ADP phosphates and the other with a single Mg^{2+} ion. Differences in coordination of phosphates and Mg^{2+} ions were noticed between both structures by means

of electron density maps. These two systems were employed as the starting models for the MD simulations. These models were solvated in a water molecules box described by the TIP3P potential while the rest of the system was treated within the AMBER force field. The flexibility of the phosphates in both systems was measured by means of root mean square fluctuations (RMSF) showing less flexibility when the system was coordinated with both Mg^{2+} ions consistent with and increased electrostatic stabilization

Mg^{2+} ions affect reaction mechanisms

Mg^{2+} ions have been also reported to be greatly important to stabilize the negative charges present during the phosphate transfer and before.[52,70] As a matter of fact, the connection between the ions function with the classification of a phosphoryl transfer reaction as an associative or dissociative has been investigated. In particular, Kamerlin and Wilkie explored PESs for the phosphate ester hydrolysis reaction with a phosphate monoanion as a model, analyzing the influence of the metal ions in the obtained mechanisms.[84] Water exchange involving phosphate monoanion ($H_2PO_4^-$) was selected as the model system. Calculations were done at Hartree-Fock (HF),[104] MP2 and DFT MPW1PW91 levels of theory using the 6-31++G(d,p) basis set. Solvation effects were simulated applying a PCM correction to DFT stationary points. The study was performed with two Mg^{2+} ions, with a single one and with the absence of metal ions. The obtained results allowed authors to conclude that a stepwise associative process was more energetically favorable than a stepwise dissociative mechanism, in presence of Mg^{2+} ions and the contrary occurred in the absence of the ions. In the case of a single Mg^{2+} ion, a concerted associative mechanism was observed displaying the largest energetic barrier. Previously, Cleland and Hengge had exposed the same conclusion in their review where the available information about nonenzymatic phosphate and sulfate transfers and the kinetic and chemical mechanisms of enzymes that catalyse similar transfers was presented.[79] They referred to phosphatases, enzymes that catalyse the hydrolysis of phosphomonoesters, explaining that an associative mechanism is favored when Mg^{2+} ions remove electronic density from phosphorus atoms, promoting the nucleophilic attack. The authors also mention that the TSs become more associative following the trend from monoesters to triesters. Nevertheless, not all the authors explain in like manner the effects caused by the ions on the phosphoryl transfer reaction mechanisms. For example, Harrison and Schulten studied the hydrolysis of ATP in aqueous

solution by means of QM/MM simulations.[105] Concretely, the computational model included the ATP molecule coordinated with an Mg^{2+} ion, inside a cubic box of water molecules. MD simulations were conducted describing the system with the CHARMM force field. The QM/MM simulations were performed using a dual grid, dual length scale method for plane-wave-based QM/MM simulations. The QM subsystem involved the ATP, the Mg^{2+} ion and 124 surrounding water molecules and it was treated using plane-wave Car-Parrinello MD method based on DFT at BLYP level. They explored the associative and dissociative alternatives and founded within the evaluation of free energies, that the dissociative process has a lower activation barrier. Moreover, authors proved that the charge transfer produced due to the presence of Mg^{2+} cation leads to lengthening of the phosphorous-oxygen bond becoming more fluctuant but compensated by an electronic stabilization with the ions. They supported that these stretching of the bond drives the reaction to a dissociative process.

In a related reaction, López-Canut et al. analyzed in their theoretical research works the different mechanisms observed in the phosphoryl transfer reactions taking place in phosphatases enzymes.[106-109] In the X-ray crystal structures employed in these works, phosphatases have two Zn^{2+} ions and one Mg^{2+} or only two Zn^{2+} ions in the active site. One of their theoretical works showed a relation between the inclusion of the two Zn^{2+} ions in the QM part and the mechanism observed.[106] Specifically, in this work the hydrolysis of the phosphate monoester, *m*-nitrobenzylphosphate (*m*-NBP^{2-}), catalysed by *Escherichia coli* alkaline phosphatase (EcAP) was studied.[106] Authors employed QM/MM hybrid potentials for evaluating free energies through the obtaining of PMFs to study the phosphoryl transfer from *m*-NBP^{2-} to the unprotonated residue, Ser102. Three different QM subsystems were defined; in the first one, only the substrate and the nucleophile were included. In the second QM model, the two Zn^{2+} metal ions were added. In the third QM subsystem the coordination shells of the ions were also incorporated. When the QM region used was the first presented, authors observed a concerted mechanism due to the overestimation of the interactions produced between the *m*-NBP^{2-} and the nucleophile. When the second model was employed the charge transfer between the reacting fragments was excessive resulting in a lower repulsion proceeding, as well, in a concerted associative pathway. However, when the calculations were done with the largest QM model, the mechanism observed was dissociative with a metaphosphate reaction intermediate since this charge transfer was reduced due to the reception of

electron density of the ions from their coordination shells, being the repulsion larger.

Different reaction mechanisms were also observed in another theoretical work of López-Canut et al. but in this case, depending on the environment of the reaction.[107] They studied the hydrolysis of the methyl *p*-nitrophenyl phosphate ($MpNPP^-$) in aqueous solution and in the active site of nucleotide pyrophosphatase/phosphodiesterase (NPP). QM/MM hybrid potentials were employed to obtain free energy profiles. In this case the phosphoryl group is transferred from the $MpNPP^-$ to the considered unprotonated Thr90 residue. In aqueous solution, the reaction was found to occur through a concerted mechanism with an associative-like TS structure. In this process, the repulsion between the negatively charged nucleophile and substrate is reduced favoring their closeness. Nevertheless, in the active site of the enzyme a concerted mechanism with a dissociative-like TS structure was observed due to the charge distribution causing by electrostatic properties of the enzymes which were able to favor this mechanism.

López-Canut et al. also observed different types of mechanisms comparing the hydrolysis of the phosphotriester paraoxon in aqueous solution and in the active site of the promiscuous *Pseudomonas diminuta* phosphotriesterase (PTE) enzyme with the hydrolysis of the phosphotriester compound O,O-diethyl *p*-chlorophenyl phosphate (DEpCPP) in the active site of the same enzyme.[109] They employed the hybrid QM/MM methodology to estimate free energies within PMFs and Kinetic Isotope Effects (KIEs). The reaction consisted in the phosphoryl transfer from the substrate (paraoxon and DEpCPP) to a hydroxide anion. Authors obtained a concerted associative mechanism of paraoxon in solution and in the active site of PTE. In contrast, the hydrolysis of DEpCPP in the enzyme proceeded by a stepwise mechanism with a pentacoordinate intermediate. The different coordination of the leaving groups corresponding to the two different substrates caused a shift of the ligands from the coordination shell of one Zn^{2+} ion to the other. The fact of having differences in the coordination shells of the metal ions justified the ability of the enzyme for catalysing hydrolysis of different substrates with different mechanisms.

Octahedral coordination sphere in Mg^{2+} ions

Another point commented by some authors about Mg^{2+} ions in kinase enzymes is the octahedral coordination sphere observed in one or both of them. This

feature was mentioned in the work of Smith and co-workers in CDK2, previously explained where the Mg^{2+} ion was observed to be coordinate with three oxygens of ATP, the side chains of Asp145, Asn132 and a water molecule.[52] These interactions assisted the reaction to take place in an in-line near attack configuration.

An octahedral coordination sphere in the metal ion was also observed in the work of Lopata et al. They studied the reaction catalysed by the wild type and two mutant forms of deoxyuridine triphosphate nucleotidohydrolase (dUTPase) performing QM/MM theoretical calculations and site-directed mutagenesis experiments.[110] This enzyme catalyse the hydrolysis of deoxyuridine triphosphate (dUTP) consisting in the transfer of a phosphate group to a catalytic water molecule where a proton is transferred from this water molecule to the Asp83 residue. QM/MM minimizations of the three systems were performed in the forward and backward directions where the system was inside a TIP3P water molecules box. In these minimizations, the quantum region contained the dUTP molecule, nearby crystallographic water molecules, the Mg^{2+} ion with its full coordination shell, the Thr81 and Ile82 backbones, the amide group of Gln113, and the side chains of different residues depending of the system (the wild type or the two mutant forms). Therefore, the wild type had 132 atoms while the two mutants 137 and 108 atoms. This QM region was described at B3LYP/6-31+G(d) level of theory while the rest of the system was treated using the CHARMM force field. In addition, short downhill QM/MM dynamics were conducted starting from the minimized wild type system and from one mutated form, near the TS. In this case, the residues side chains included in QM region of the two systems were a bit different as well, resulting in 121 atoms for the wild type system and 95 atoms for the mutated enzyme. A one-step mechanism within a dissociative TS was observed in the three enzymatic systems. The metal ion coordination was analysed during the reaction proving that an octahedral coordination was essential to drive the reaction towards products since it stabilized the transition state structures. Authors also mentioned the fact that the Mg^{2+} to Ca^{2+} substitution effects are normally inhibitory in kinases although there are some exceptions. The work of Bastidas et al.[102] pointed out previously showed through their MD simulations and their experimental results an octahedral coordination for both Mg^{2+} ions, emphasizing in the strong binding character of specially one since it recruits a water molecule after the phosphate transfer in order to keep on with its coordination geometry. In this way, the metal ions

help to catalyse the reaction and to stabilize the transition state. An example of this coordination sphere characteristic in Mg^{2+} ions is presented in Figure 4.2.

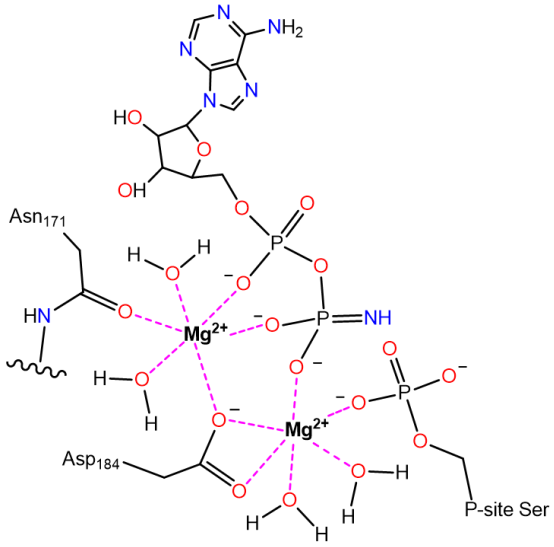


Figure 4.2. Octahedral coordination observed in both Mg^{2+} cations where the phosphoryl group has been transferred from AMP-PNP to a Ser side chain substrate in PKA. Figure adapted from reference.[102]

4.3. Molecular mechanism of PTS-dependent DHAK from *Escherichia coli*

As it has been pointed out in the beginning of this section, DHAKs are divided into two classes, those using the PEP carbohydrate: PTS, to provide the phosphoryl group, in most bacteria or those using ATP as the phosphoryl donor in animals, plants and some bacteria.[69,70,74-78] This section is devoted to DHAK from *Escherichia coli* (*E. coli*) which is considered a prototype enzyme of the PTS-dependent family.

The PTS is a multicomponent system in bacteria implicated in the consumption and concomitant phosphorylation of numerous carbohydrates.[111,112] It was discovered in *E. coli* by Kunding, Ghosh and Roseman as a system that uses PEP for the phosphorylation of hexoses.[113] The two general phosphoryl transfer proteins of the PTS are Enzyme I (EI) and the heat-stable, histidine-phosphorylatable protein (Hpr).[75,111,112]

DHAK from *E. coli* is a heterotetramer complex containing a central DhaK dimer and two molecules of DhaL. Each DhaL subunit interacts with only one DhaK subunit of the dimer. An ADP molecule and two magnesium ions are bounded to DhaL, whereas DhaK contains a glycerol molecule covalently bound to His218.[70] In Figure 4.3 is shown the complex DhaK-DhaL. This enzyme also contains an additional subunit, DhaM. This subunit has three domains, two of them homologous to the PTS proteins which are EI and HPr, as pointed out before. Concretely, the N-terminal domain of DhaM has the same fold as the IIA domain of the PTS transporter for mannose of *E. coli*. The middle domain is similar to HPr and the C-terminal is similar to the N-terminal domain of EI of the PTS.

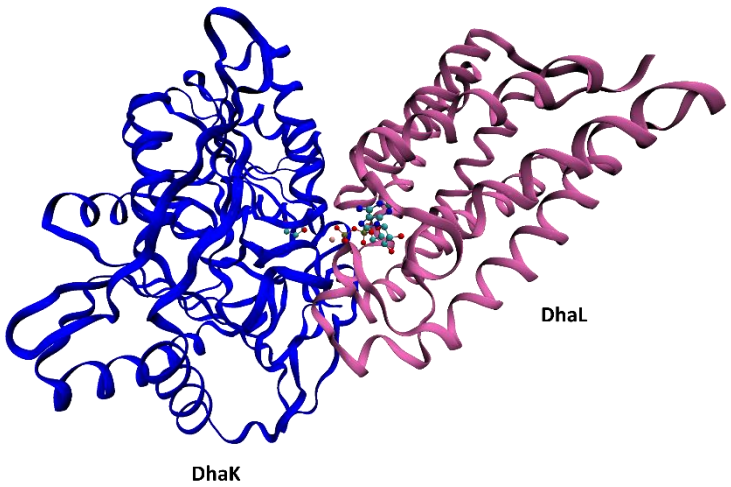
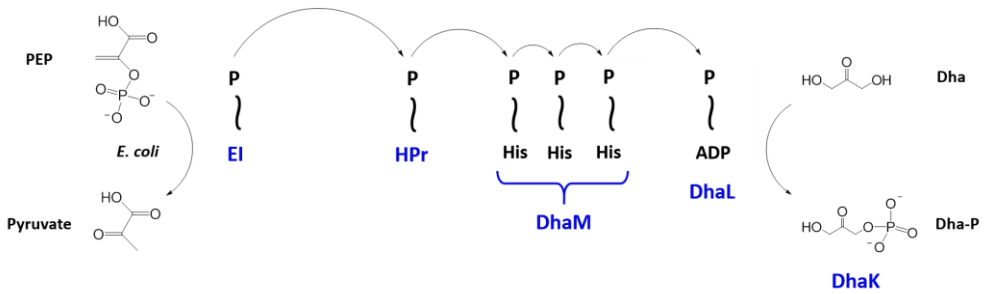


Figure 4.3. Cartoon representation of the DhaK-DhaL subunits from the PDB code crystal structure 3PNL. DhaL molecule is shown in mauve and DhaK subunit is colored in blue. The glycerol molecule placed in DhaK and the ADP molecule and magnesium ions bound to DhaL are shown as balls and sticks model.

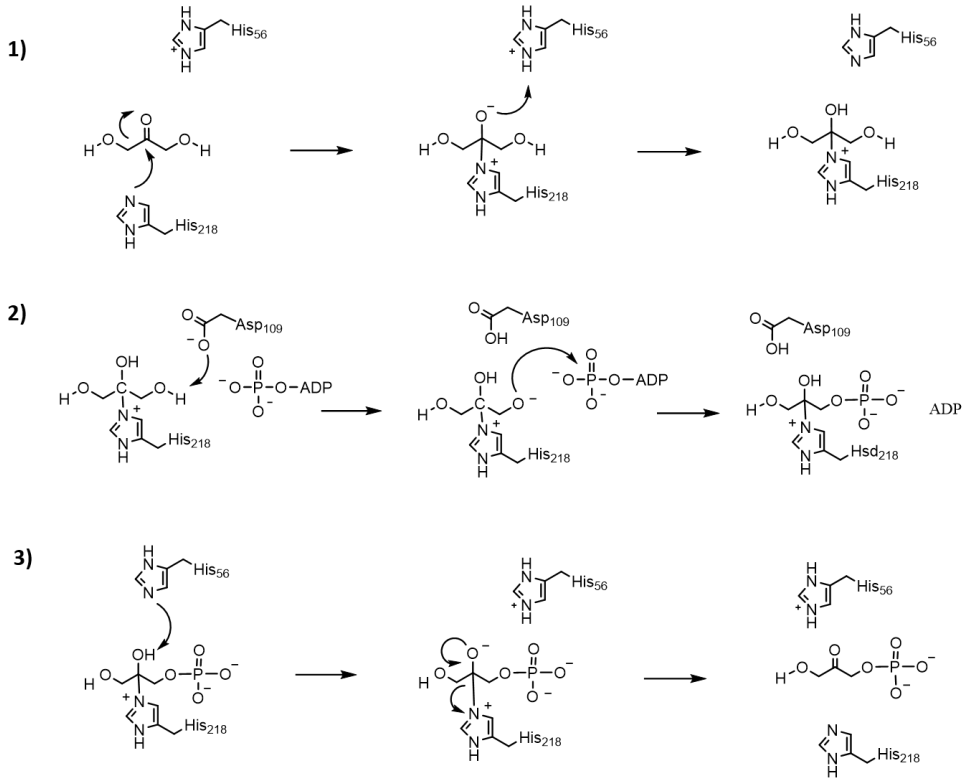
PEP phosphorylates three times DhaM through an EI- and HPr- dependent reaction while DhaK and DhaL are not phosphorylated.[74,75,78,114] Thus, in the first place, EI catalyses the phosphoryl transfer from PEP to HPr. Subsequently, this small phospho-carrier protein serves as phosphoryl donor to DhaM. Phosphorylated DhaM forms a complex with the ADP-bound DhaL subunit and transfer the phosphate to ADP turning it into ATP. Finally, phosphate is transferred from ATP to Dha yielding the product of the reaction,

Dha-P and ADP.[69,70,78,114] All this procedure of the phosphate transfer from PEP to Dha is schematically shown in Scheme 4.6.



Scheme 4.6. Schematic procedure of the PTS system. Phosphoryl group is transferred in a sequential way from PEP through EI to HPr, then to DhaM, subsequently to ADP bounded to DhaL, and finally to Dha placed in the DhaK subunit. Arrows indicate the phosphate transfer reactions. DhaM is phosphorylated three times in three histidine phosphorylation sites.[115] Figure adapted from reference[75].

The molecular mechanism of the phosphoryl transfer from ATP to Dha has been proposed based in the crystal structure of the *E. coli* DhaK-DhaL complex bound to ADP and Dha (3PNL) determined by Shi and co-workers.[70] Analysis of experimental results allowed them to present a general catalytic mechanism for DHAK in which residues His56, His218 and Asp109 would be significantly involved in the reaction progress. In particular, the authors divided the whole reaction in three stages, which are presented in Scheme 4.7: 1) binding of Dha to enzyme through an hemiaminal bond with His218. In this stage, His 56 acts as an acid donating its proton to Dha substrate. 2) Phosphate transfer from ATP to Dha where Asp109 acts as a base in order to prepare the Dha for the nucleophilic attack. 3) Release of the reaction product, Dha-P, from the enzyme.[70]



Scheme 4.7. Catalytic mechanism of the phosphoryl transfer from ATP to Dha in DHAK from *E. coli* proposed by Shi and co-workers. 1) Dha binding to the enzyme and proton transfer from His56 to the Dha substrate. 2) Abstraction of the Dha proton by Asp109 followed by the phosphoryl transfer to Dha. 3) Proton transfer from Dha to His56 and cleavage of the His218 bond with the consequent release of the Dha-P product. Figure adapted from reference.[70]

As noted in Scheme 4.7 this mechanism would classify as an *asp-assisted mechanism* (see Section 4.1) therefore, the possibility of a *substrate-assisted mechanism* has also considered in this PhD Thesis. In fact, both mechanisms have been explored by means of theoretical methods in DHAK from *E. coli* (see chapter 7.2) carrying out a comparative energetic and structural analysis of the two mechanisms. Through this research work it has been justify whether the mechanism proposed by Shi et al. is the most favorable one in this enzyme.

4.4. ATP-dependent DHAK from *Citrobacter freundii*. Inorganic polyphosphate as phosphoryl donor

This section in dedicated to a representative DHAK of the ATP-dependent family which is the DHAK from *Citrobacter Freundii (C. freundii)*. This enzyme is a homodimer and each subunit is formed by two domains.[74,77,116] The DhaK-domain is where the Dha binding site is located and in the DhaL-domain is founded the ATP binding site coordinated with the two Mg^{2+} ions. In the dimer, the subunits are arranged in an anti-parallel way, therefore, the DhaK-domain of one subunit is faced with the DhaL-domain of the other subunit. The DhaK-domain has the same structure as the DhaK subunit of *E. coli* DHAK. The ATP binding domain exhibits a novel alpha-fold. It is formed by eight amphipathic and antiparallel alpha-helix organised following an up-and-down geometry. They are forming a circle resulting in a barrel confining a cavity where is located a lipid.[73,116]

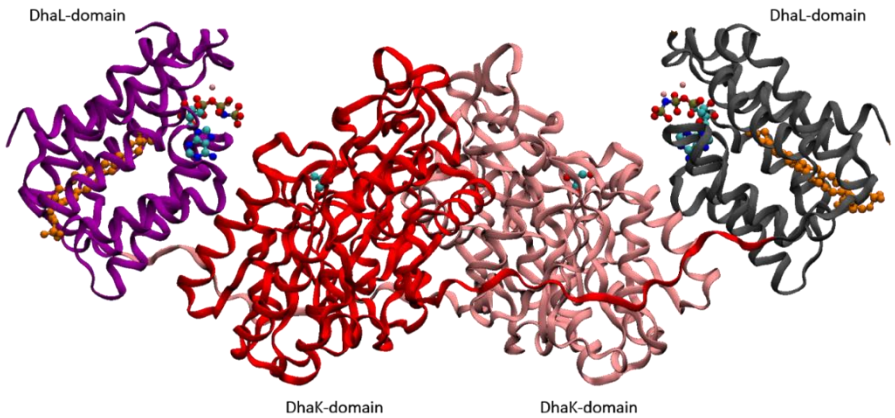
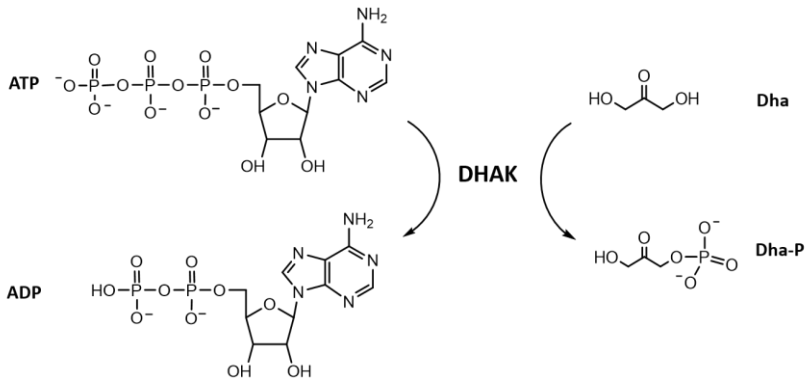


Figure 4.4. Cartoon representation of the kinase homodimer obtained by superposition of PDB code crystal structures 1UN8 and 1UN9. The DhaK-domains are colored in pink and red, and the DhaL-domains are colored in purple and grey. Shown as ball and sticks model are: Dha, phosphoaminophosphonic acid-adelynate ester (ANP), magnesium ions and phospholipid in orange.

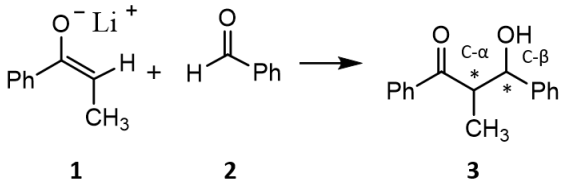
In the ATP-dependent kinase, ATP/ADP exchange is fast as is schematically shown in Scheme 4.8.[74]



Scheme 4.8. Phosphoryl transfer reaction from ATP to Dha in DHAK from *C. freundii*. Figure adapted from reference[75].

- Aldolase activity

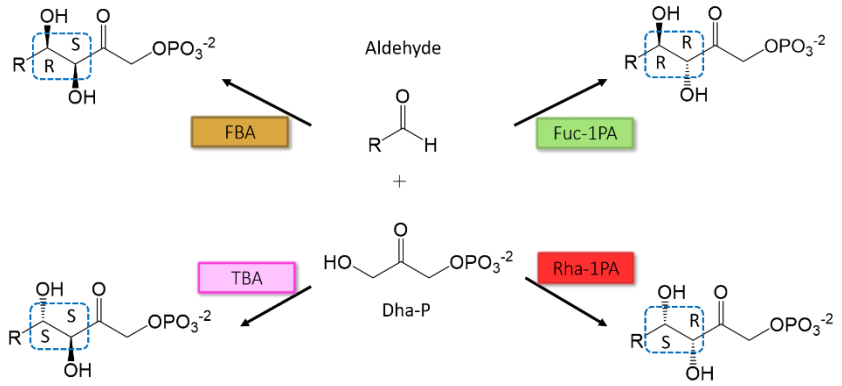
Aldol addition reaction has long been recognized as one of the most useful tools that has the synthetic chemist for the creation of new C–C bonds.[73,117-119] An example of an aldol reaction is displayed in Scheme 4.9:



Scheme 4.9. Example of an aldol addition reaction; a ketone enolate (1), acting as a nucleophile, is added to the electrophilic carbon of an aldehyde (2), forming a new carbon-carbon bond to give the aldol product.

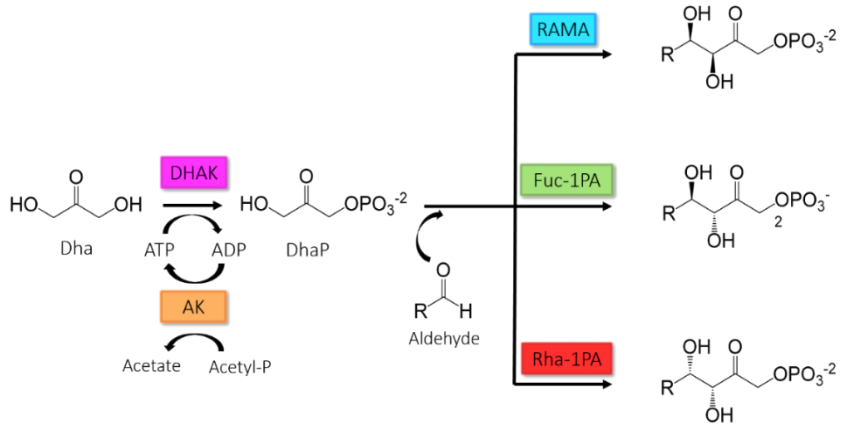
There are enzymes in nature that catalyse these reactions; the aldolases. The main group of aldolases is the one that uses Dha-P as substrate donor; Dha-P dependent aldolases. They are among the most important biocatalysts for C-C bond formation.[73,120-124]

Aldol reactions catalysed by Dha-P dependent aldolases consist in the formation of two new stereocenters with a stereochemistry controlled by the enzymes. Therefore, from two given substrates it is possible to obtain the four diastereoisomers using four different aldolases which is a powerful synthetic advantage, as shown in Scheme 4.10.[73,124-126]



Scheme 4.10. Aldol reactions catalysed by the four aldolases; fuculose-1-phosphate aldolase (Fuc-1PA), rhamnulose-1-phosphate aldolase (Rha-1PA), fructose-1,6-biphosphate aldolase (FBA) and tagatose-1,6-biphosphate aldolase (TBA). Figure adapted from reference[121,125].

However, their major disadvantage is precisely their strict specificity for the expensive and unstable Dha-P.[73,124,126,127] This molecule is overpriced to be used in large-scale synthesis and is labile at neutral and basic pH values causing the decrease of its effective concentration with time in the enzymatic reaction media.[123] Therefore, an efficient method of Dha-P preparation is still essential.[73,128] Following this requirement, the group of Prof. Eduardo García-Junceda of the Institute of the General Organic Chemistry in Madrid have elaborated a straightforward multi-enzyme system for one-pot C–C bond formation catalysed by DHAP-dependent aldolases, based in the use of the ATP-dependent DHAK from *C. freundii* for in situ Dha-P formation.[73,120,122,123] In this system the ATP is regenerated by catalysis of acetate kinase (AK). The fusion protein preserve both kinase and aldolase activity with a very high catalytic efficiency.[73,120,122,123] This system is presented in the Scheme 4.11.



Scheme 4.11. Operating scheme of the multi-enzyme system for C-C bond generation catalysed by DHAP-dependent aldolases. Dha substrate is phosphorylated *in situ* by DHAK from *C. freundii* while the ATP is regenerated by AK. In this scheme; RAMA: rabbit muscle aldolase. Figure adapted from reference[73].

Inorganic polyphosphate as a phosphoryl donor. Mutation in DHAK from *C. freundii*

Despite of the benefits of the multi-enzyme described above there is one disadvantage; the requirement of an ATP regeneration system since the direct addition of ATP is often problematic due to the formation of inhibitory products such as ADP or AMP.[129,130] Furthermore, high concentrations of ATP inhibit some enzymes such as phosphofructokinase.[129]

Several ATP regeneration systems have been created through the employment of biological agents including whole cells, organelles and enzymatic systems.[130] Some of these enzymatic systems have been produced using several combinations of phosphate donors and enzymes,[131,132] such as acetyl phosphate (AcP) and acetate kinase, PEP with pyruvate kinase, and creatine phosphate (PC) combined with creatine kinase.

ATP regeneration was applied to enzymatic synthesis of guanosine monophosphate (GMP), cytidine diphosphate-choline (CDP-choline) and other materials in the industry. These systems are suitable for particular applications but there are problems regarding to the expensive prices of chemicals and the lack of accessibility to a method for regenerating ATP from AMP.[131] Therefore, it is necessary to find a phosphoryl donor cheaper than ATP and, more important, that the final product does not inhibit the kinase. A

very good candidate is the poly-P.[3,73,130] Inorganic poly-P is a linear polymer of up tens to hundreds of phosphates linked by high-energy phosphoanhydride bonds.[73,133,134]

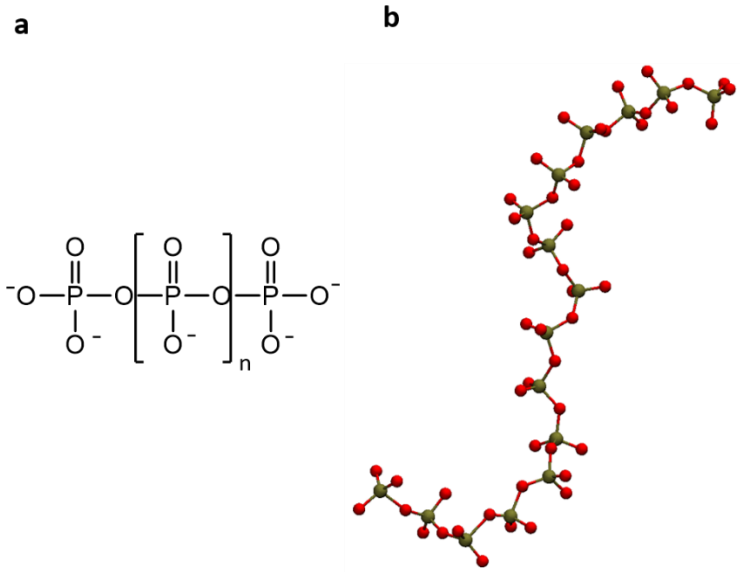


Figure 4.5. Scheme of inorganic polyphosphate (a) and 3D picture of inorganic polyphosphate formed by sixteen phosphate residues (b).

Regarding to the price of this compound, a commercial form of poly-P which costs $9/lb is able to provide ATP equivalents that would cost over $2,000/lb separately, while there are other phosphagens capable of regenerating ATP that cost more than ATP, such as PEP and phosphocreatine.[129] In addition, a huge quantity of poly-P is regularly produced as sodium hexametaphosphate (about 13 to 18 residues) for industrial uses such as food additives which also makes poly-P inexpensive compared to other phosphoryl donors.[73]
Nevertheless, the wild type DHAK from *C. freundii* does not show catalytic activity with the polymer. Therefore, García-Junceda and co-workers initiated an experimental directed evolution program with the aim of transforming the specificity of the phosphoryl donor in this enzyme from the ATP to the poly-P. They found that sixteen mutant clones exhibited an activity with poly-P as phosphoryl donor statistically relevant and the most active mutant presented a single mutation, Glu526Lys, located in a flexible loop. Our group have collaborated with the experimental group of García-Junceda by means of computational calculations (see Chapter 7.3) in order to analyse the properties

of the mutant form of the enzyme and thus, provide information about the origin of its catalytic activity observed in the experimental results. Therefore, theoretical calculations have focused on the wild type and on the most active mutant allowing to perform a comparative study of the evolution of the interactions between the protein and the poly-P in both systems.

5. Computational methods

5.1. QM/MM hybrid potentials

To accurately model the electronic-structure problem involving bond-making and bond-breaking processes in chemical reactions, charge transfer or electronic excitation, the employment of Quantum Mechanical (QM) methods are required. However, the description of all the particles of an entire biomolecular system through the QM methodology is unworkable since these theoretical methods require complex mathematical models and thus need a large number of computational resources.[135] This problem can be overcome by using Molecular Mechanical (MM) methods also known as force fields, which are based on parameters derived from experimental data. In these methods the electronic motions are ignored and the calculation of the energy of a system is a function of the nuclear positions which allows to apply this methodology in a region containing a high number of atoms, with a low computational cost. Nevertheless, the MM methods are not adequate to describe chemical reactions.[136,137] Therefore, the combination of both methodologies that results in the Quantum Mechanical/Molecular Mechanical (QM/MM) hybrid methods, first proposed by Arieh Warshel, Michael Levitt and Martin Karplus in the 1970s,[138-141] is the most suitable computational strategy for the study of enzymatic reactions. It allows the inclusion of environment effects over the reaction mechanism in the computational simulations, which is fundamental when enzymatic reactions are the purpose of the study. Different types of QM/MM methods are available and their use is applied to developments in drug metabolism and in drug and biocatalyst design, having an impact not only in enzymology but also in other subjects of biochemistry and medicinal chemistry.[142] The importance of QM/MM potentials in comprehension of biological systems was confirmed in 2013, with the award of the Nobel Prize for Chemistry to Warshel, Levitt and Karplus.

First of all, in order to use this hybrid methodology, the QM and the MM regions of the system should be selected as shown in Figure 5.1.

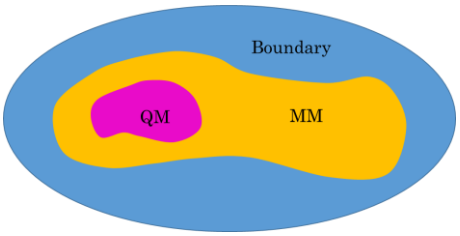


Figure 5.1. Schematic representation of the different regions of a QM/MM model. In pink is highlighted the QM region, where the chemical reaction takes place, in orange is displayed the rest of the system and the colour blue represents the boundary conditions.

The QM region should contain, at least, the atoms which participate directly in the formation or breaking of the covalent bonds. Usually, it includes the substrate, the cofactor (if the system need it), may also include solvent fragments and the amino acids of the active site of the enzyme which are involved directly in the reaction or by means of covalent bond interactions or electron density transfer with the substrate or cofactor.
The MM region contains the rest of the atoms in the system, that is, the rest of the amino acids of the enzyme, the solvent molecules and counterions.
Finally, the boundary conditions are a collection of restrictions applied to the system due to the impossibility of simulating an infinite system, as it will be explained later on.
To be able to compute the energy and the forces of each atom in the system, an effective Hamiltonian, $\hat{H}_{eff}$, is constructed and the time independent Schrödinger equation is solved for the energy of the system, E, together with the wavefunction, Ψ , for the electrons of the atoms in the QM part:[140]

$$\hat{H}_{eff}\Psi(r, R_\alpha, R_M) = E(R_\alpha, R_M)\Psi(r, R_\alpha, R_M) \qquad (5.1)$$

Ψ depends on the electrons' coordinates (r), and on the position of both the quantum and classical nuclei (R_α and R_M respectively). In this equation the Born-Oppenheimer approximation[143] has been applied where a differentiation in the treatment of the nuclear and electronic motion is assumed since the nuclei has a greater mass than the electrons. Thus, their movements can be considered independently.
This wavefunction describes the electronic density of the quantum region under the influence of the environment. The energy is calculated for each

position for the QM nuclei and MM atoms, using the following equation which is solved by means of a self-consistent process:

$$E = \frac{\langle \Psi | \hat{H}_{eff} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{5.2}$$

In order to construct the effective Hamiltonian of the system, it should be considered that QM/MM schemes can be classified, regarding to the nature of the definition of the total energy, into two extreme schemes;[144,145] the subtractive and the additive which is the one applied in this PhD Thesis.
The total Hamiltonian for a subtractive scheme would be:

$$\hat{H}_{eff} = \hat{H}_{QM(QM,Link)} + \hat{H}_{MM(QM,MM)} - \hat{H}_{MM(QM,Link)} + \hat{H}_{BC} \tag{5.3}$$

Equation (5.3) describes a scheme with link atoms (see later on in this section) where the QM region is separated and threated at a QM level ($\hat{H}_{QM(QM,Link)}$). The MM methodology is applied to both, the QM and the MM region ($\hat{H}_{MM(QM,MM)}$) and subsequently is subtracted the classical energy corresponding to the QM region ($\hat{H}_{MM(QM,Link)}$). The main advantage is that no explicit QM/MM coupling terms are required. The $\hat{H}_{BC}$ term will be developed then, in this section. Examples of a subtractive schemes are the integrated molecular orbital/molecular mechanics (IMOMM) method[146] or the n-layered integrated molecular orbital and molecular mechanics (ONIOM) approach.[147-149]
Regarding to the additive QM/MM scheme, the effective Hamiltonian, $\hat{H}_{eff}$, consists of:

$$\hat{H}_{eff} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} + \hat{H}_{BC} \tag{5.4}$$

Each of its components are developed below:

$\hat{H}_{QM}$ describes the electrons, nuclei and interactions among the atoms of the QM part. It depends on the quantum method used and corresponds to an electronic Hamiltonian of the subsystem in vacuum:

$$\hat{H}_{QM} = \hat{K}_i + \hat{V}_{ij} + \hat{V}_{i\alpha} + \hat{V}_{\alpha\beta} \tag{5.5}$$

$$\widehat{H}_{QM} = -\frac{1}{2}\sum_{i}\nabla_i^2 + \sum_{ij}\frac{1}{r_{ij}} - \sum_{i\alpha}\frac{Z_\alpha}{r_{ia}} + \sum_{i\alpha}\frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \tag{5.6}$$

Equations (5.5) and (5.6) represent the sum of the kinetic energy of the electrons and the potential energy for the electron-electron repulsion, nucleus-electron attraction and nucleus-nucleus repulsion. In this equation i, j are electronic coordinates, α, β are nuclear coordinates, r is the electron-electron or nucleus-electron distance, R is the nucleus-nucleus distance, Z is the nuclear charge and ∇ is the kinetic energy or Laplacian operator.

$\widehat{H}_{MM}$ describes the interactions between the atoms defined by means of molecular mechanics. As stated in the beginning of the section, the MM energy is calculated by means of a force field. This term can take out of the integral (equation 5.2), since it only depends on the MM atoms and not on the electronic coordinates of the QM atoms. It is depicted as:

$$\widehat{H}_{MM} = E_{MM} \tag{5.7}$$

$$E_{MM} = E_{bonded} + E_{non-bonded} \tag{5.8}$$

Where E_{bonded} is:

$$\begin{aligned} E_{bonded} &= E_{bonds} + E_{angles} + E_{dihedrals} \\ E_{bonded} &= \sum_{bonds}\frac{1}{2}k_b(b-b_0)^2 + \sum_{angles}\frac{1}{2}k_\theta(\theta-\theta_0)^2 + \\ &\quad + \sum_{dihedrals}\frac{1}{2}V_n[1+cos(n\emptyset-\delta)] \end{aligned} \tag{5.9}$$

In the first term of the equation (5.9), k_b is the force constant for the bond, b is the actual bond length and b_0 is its equilibrium distance. In the second term k_θ corresponds to the force constant for the angle, θ is the actual angle and θ_0 is the equilibrium value for the angle. Lastly, in the third term V_n means the force constant of the dihedral angle, n is its periodicity, $\emptyset$ corresponds to the dihedral angle and δ to its phase.
The $E_{non-bonded}$ term of the equation (5.8) is represented as:

$$E_{non-bonded} = E_{electrostatic} + E_{van\ der\ Waals}$$

$$E_{non-bonded} = \frac{1}{4\pi\epsilon_0\epsilon} \sum_{ij\ pairs} \frac{q_i q_j}{r_{ij}} + \sum_{ij\ pairs} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \qquad (5.10)$$

In the first term of this equation, q_i and q_j are the fractional charges of atoms i and j, r_{ij} is the distance between the two particles, $1/4\pi\epsilon_0$ is the standard term when electrostatic interactions are calculated and ϵ is the dielectric constant. In the second term A_{ij} and B_{ij} are positive constants whose values depend on the types of atoms i and j.

As can be seen, the atoms are represented by punctual charges with their corresponding van der Waals parameters, in order to calculate the interactions between those that are not connected by a bond. The short distance interactions are evaluated using terms related to bond distances, angles and dihedrals that describe the connectivity of the molecule. Therefore, the set of energetic terms and parameters describing the system constitute a force field. The generic form of the bond and angle terms is usually quadratic (harmonic potential), while for the torsion term, developments of Fourier series are employed and for electrostatic and van der Waals terms, Coulomb and Lennard-Jones expressions are normally employed respectively.[150] Non-bonded terms should calculate for all atom pairs but those which are separated by 3 or 4 bonds depending on the force field thus, eliminating from the non-bonded calculation the contribution of the directly bonded atoms (1-2 interactions) and those which are separated by two bonds (1-3 interactions). Contributions of atoms separated by three bonds (1-4 interactions) are calculated or not depending on the force-field employed, although this interaction is usually scaled multiplying by a factor (see Figure 5.2).[150] The force field and the parametrization employed for the calculations in this PhD Thesis are based in CHARMM22[151-153] and OPLS-AA[154] force fields to describe the proteins, and the TIP3P[155] force field for the water molecules.

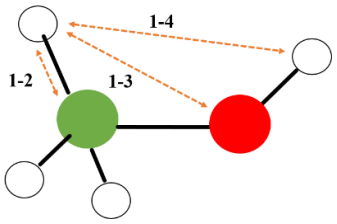


Figure 5.2. Representation of the 1-2, 1-3 and 1-4 interactions in the methanol molecule (CH_3OH).

$\widehat{H}_{QM/MM}$ describes the interactions between the QM and MM atoms and it can be defined as the contribution of three terms: electrostatic, van der Waals and polarization.

$$\widehat{H}_{QM/MM} = \widehat{H}^{elec}_{QM/MM} + \widehat{H}^{vdW}_{QM/MM} + \widehat{H}^{pol}_{QM/MM} \qquad (5.11)$$

$\widehat{H}^{pol}_{QM/MM}$ term is usually removed from the equation in most implementations since the polarization of the MM subsystem would result in solving the total Hamiltonian in a self-consistent manner due to the change of the quantum wavefunction because of the charges variation.
Removing this polarization term and due to the fact that the MM atoms are represented by punctual charges and van der Waals parameters as pointed out above, this Hamiltonian is expressed as:[150]

$$\widehat{H}_{QM/MM} = -\sum_{iM}\frac{q_M}{r_{iM}} + \sum_{\alpha M}\frac{Z_\alpha q_M}{R_{\alpha M}} + \sum\left\{\frac{A_{\alpha M}}{R^{12}_{\alpha M}} - \frac{B_{\alpha M}}{R^{6}_{\alpha M}}\right\} \qquad (5.12)$$

Where q_M is the charge of the classical atoms, Z_α is the nuclear charge of the quantum atoms, the subscript i refers to the electrons of the QM atoms, the subscript α is the nuclei of the QM atoms and the subscript M refers to the MM atoms.
The first term of the equation (5.12) represents the electrostatic interactions between the MM atoms and the electrons of the QM part. The second term is the electrostatic interactions between the MM atoms and the nuclei of the QM atoms. The last term describes the van der Waals interactions between the MM and QM atoms, being A and B the parameters of a 6-12 Lennard-Jones potential that models these interactions, presented also in the equation (5.10). These parameters are transferable among systems since they are exclusively

dependent on the type of atoms involved.[156] In this equation, just the first term contains the electrons' coordinates, so, it must be incorporated in the self consistent field (SCF) procedure with que quantum Hamiltonian. The remaining terms are constant for a given set of atomic coordinates, like the MM Hamiltonian ($\widehat{H}_{MM}$).

$\widehat{H}_{BC}$ refers to the attempt of simulating a condensed phase (in this case, a solvent) which is infinite at an atomic level, with a finite number of molecules and thus, the environment of the system is well imitated. There are different approximations to model the environment of a system but the most widely used, which is the employed in this PhD Thesis, is the method of periodic boundary conditions (PBC). Following this approach, the system is introduced in a solvent box which is replicated in an infinite way in the three dimensions of the space, as indicated in Figure 5.3. Therefore, system molecules situated in the unit cell interacts with replicated images being able to simulate the environment formed by solution through long-distance interactions. Obviously, the finite system should have a regular shape in order to fill the space when it is replicated, what is achieved using very regular boxes such as cubic or orthorhombic ones.

The employment of PBC implies the use of the *minimum image convention* which assumes that each particle only interact with its nearest copy in the system when the non-bonding energy is calculated.[150] This fact makes the method cheaper and easier to implement, but it means that the cutoff distance for truncation of the non-bonding interactions must be less than or equal to half the length of the side of the periodic box.

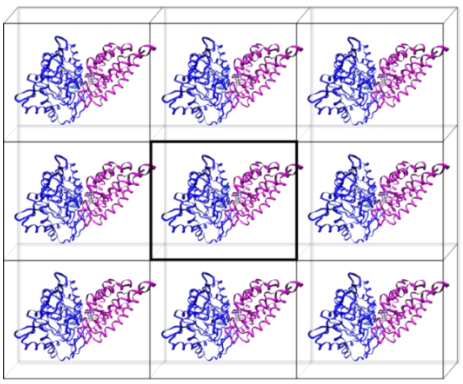


Figure 5.3. Schematic representation of the employment of PBC where the central square cell highlighted is replicated in the three dimensions of space.

Cutoff methods evaluate the non-bonding energy over all the atom pairs of the system with a reasonable computational cost and without losing precision in results. The strategy consists in set up truncation distances from which the non-bonded interactions are ignored. Then, only pair interactions inside the cutoff radio defined will be considered. There are different truncation functions that can be applied.[150] The one used in this PhD Thesis is a *Switch function*; it computes all the interactions inside a radius (r_{on}), from which it is smoothed to zero at the outer cutoff (r_{off}). Thus, all pairs with distances greater than r_{off} will not include in the interaction calculations:

$$S(r) = \begin{cases} 1 & r \leq r_{on} \\ \dfrac{\left(r_{off}^2 - r^2\right)^2 \left(r_{off}^2 + 2r^2 - 3r_{on}\right)}{\left(r_{off}^2 - r_{on}^2\right)^3} & r_{on} < r \leq r_{off} \\ 0 & r > r_{off} \end{cases} \qquad (5.13)$$

Moreover, in most of the molecular systems under study, a covalent bond has to be cut when the MM and QM regions are separated which implies that a particular molecule must be divided between the QM and the MM region. As pointed out in the beginning of this section, in the study of the reactions in biological systems there are frequently some aminoacids in the active site which are involved actively in the reaction, thus, they participate in the breaking and forming of covalent bonds. Therefore, part of these residues should be in the QM region as well as residues implicated in charge transfer with the reactant system. Besides deciding where cutting will occur, it is required to do approximations in order to satisfy the valences of the QM atoms. Among the methods for performing the partition of a molecule, the one applied in this PhD Thesis in the ***link atom methodology***[140,157] in which the MM part of the bond is replaced by an atom, usually a hydrogen atom since the C and H electronegativities are similar (see Figure 5.4). This link atom is used to saturate the valences of the truncated bond and is invisible to the MM region because the interactions between the link atom and the MM atoms are not considered.

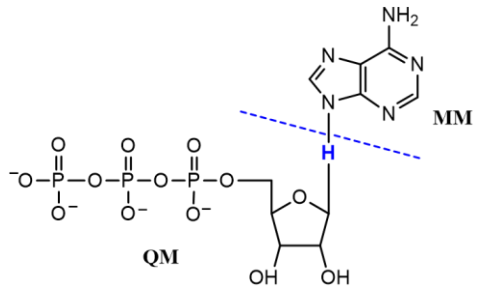


Figure 5.4. Representation of the link atom methodology for the frontier atoms treatment in the ATP molecule. The H atom coloured in blue is the link atom.

Usually, the selected boundary atom is non-polar to allow a correct separation of the two electrons of the bond: one will stay in the QM region and the other will be removed in the MM region.

Finally, the total energy of the system can be expressed as the sum of the values of each term of the Hamiltonian:

$$E = E_{QM} + E_{MM} + E_{QM/MM} + E_{BC} \tag{5.14}$$

Applying the integral of the equation (5.2) and considering that E_{BC} can be divided in QM and MM, the energy expression results in:

$$E = \frac{\langle \Psi | \hat{H}_{QM} + \hat{H}_{QM/MM} + \hat{H}_{BCQM} | \Psi \rangle}{\langle \Psi | \Psi \rangle} + E_{MM} + E_{BCMM} \tag{5.15}$$

The forces on the QM nuclei, F_α, and MM atoms, F_M, are obtained differentiating the energy equation (5.15) with respect to the coordinates of the QM nuclei and MM atoms respectively.[140]

$$F_\alpha = -\frac{\partial E}{\partial R_\alpha} \tag{5.16}$$

$$F_M = -\frac{\partial E}{\partial R_M} \tag{5.17}$$

5.1.1. QM methods

As pointed out previously, QM methods are employed to describe the electronic rearrangement that occurs in chemical reactions, when breaking and forming bonds takes place. The Born-Oppenheimer approximation allows to ignore the electron-nuclear correlation, however, the correlation between electrons is a challenge when poly-electronic systems are the object of study. Therefore, the aim of QM methods is to solve the Schrödinger equation getting an approximate wavefunction and energy for the electronic problem. The complex mathematical models required by these methods restricts their use to a reduced number of atoms that corresponds to the QM region selected in Figure 5.1. Depending on the QM method employed the methodological approximations to achieve this goal will be different and thus, the accuracy and the expense of the calculations. Therefore, depending on these approximations, QM methods can be generally divided into three types: *ab initio*, density functional theory (DFT) and semiempirical approaches.[10]

- *Ab initio methods; Hartree-Fock theory*

The *ab initio* QM methods make use of the Hartree-Fock (HF) theory,[158] which was the first approach to solve the electronic problem. The HF approach is also called Self Consistent Theory since it follows an iterative process to solve the Schrödinger equation until reaching the autoconsistency or convergency. The group of *ab initio* methods concern quantum mechanical calculations obtained exclusively from theoretical principles, without including experimental data. HF methodology basically assumes that the electrons are distinguishable and independent from each other which is clearly an approximation since the electrons are indistinguishable. In addition, it considers that the electrons move in the field of the average potential originated by the rest of the electrons which is also an approximation because the movement of the electrons depends on the instantaneous position of the other electrons, that is, their movements are correlated. However, there are different methods which introduce the correlation energy: the conventional methods, originated from the HF function, called post-Hartree-Fock methods, and the non-conventional methods that include the electronic correlation in an alternative way, which are based in the Density Functional Theory (DFT).[159]

- *Methods based in Density Functional Theory*

DFT was developed by Pierre Hohenberg and Walter Kohn in 1964.[93] They affirmed through the Hohenberg-Kohn theorem that the energy of the ground electronic state of a system can be determined if its electronic density is known. This theorem was improved by Kohn and Sham who postulated that a system with 2N electrons without interacting and described by molecular orbitals would present the same electronic density than the real system, with interactions, if the correlation and exchange functional were exactly known. Since these methods include correlation effects they are more accurate than HF with a computational cost relatively cheaper. In addition, DFT methods have been employed in many computational studies of phosphoryl transfer reactions in gas phase, with implicit solvent models or in the active site of enzymatic systems.[87,89,160-163] Among the methods that follows this theory, the functional *Becke three-parameter Lee-Yang-Parr* (B3LYP)[94,95] has been the one selected to use in this PhD Thesis. This functional has been employed in a variety of systems,[164-167] including works of phosphate hydrolysis[81,107,168] and in phosphoryl transfer reactions in kinases[48,89,169] giving activation energies in agreement with results of experimental measurements. It has shown to render phosphate compounds with good geometries and energies in a feasible computational time.

- *Semiempirical methods*

Semiempirical methods have been developed with the main objective of studying systems of chemical interest with a much lower computational cost. Therefore, they are employed frequently to study big systems such as biomolecules. These methods only consider the valence electrons since these electrons indicate the chemical properties of the elements. The intern electrons of the atoms are integrated inside a *core* with the nucleus. Another approximation included in semiempirical methods is the *Zero Differential Overlap* which implies that a lot of the monoelectronic and most of the bielectronic integrals are annulated. The rest of the integrals are made equal to parameters which result from least square adjustments with the aim of reproducing values of particular magnitudes for a set of molecules.[159] These parameters are introduced to compensate the fact that some terms of the Hamiltonian are not computed explicitly and can be derived from experimental properties or from *ab initio* calculations on model systems. The semiempirical methods used in this PhD Thesis are: the Austin Model 1 (AM1),[170] the modified semiempirical Hamiltonian AM1/d-Prot (AM1d),[100] the parametric

method number 3 (PM3)[171] and the parametric method number 6 (PM6).[172] Semiempirical Hamiltonian AM1d has been developed to model phosphoryl transfer reactions since it includes d-extension for the phosphorus atom. It has modified parameters for oxygen and hydrogen atoms and the rest of the atoms are described at the AM1 level. AM1d has been employed in the study of phosphoryl transfer reactions providing results in very good accordance with DFT calculations and with experimental results.[108,109,173,174] The PM3 Hamiltonian is normally 3-4 orders of magnitude faster than DFT methods and it has proved to produce a huge energy stabilization in phosphorane compounds resulting from the employment of a minimal valence basis.[100] Therefore, it has been employed to properly model phosphorous and phosphate groups.[65,175-177] In spite of its inaccuracy to model phosphoryl transfer reactions due to the lack of d orbitals, its low computational cost allows to perform statistical simulations to rapidly explore free energy surfaces and to set up in a practicable manner an optimal computational protocol.[177] PM6 is a more complete optimization parameter resulting from several changes to the core-core approximations into the Neglect of Diatomic Differential Overlap (NDDO) methodology. The addition of d orbitals to the main-group elements and the introduction of diatomic parameters were the most important modifications.[178]

In this PhD Thesis, it has been made a comparative study among the semiempirical methods AM1d, PM3 and PM6 to analyse the phosphoryl transfer reaction from ATP to Dha in aqueous solution. It has been proved that the employment of different Hamiltonians results in different reaction mechanisms and different values of activation and reaction energies. The study of the reaction in aqueous solution presented low free energy barriers at PM3 semiempirical level as well as good reaction product energies and suitable geometries for all the stationary points. When this phosphoryl transfer reaction was explored in the active site of the DHAK enzyme it was founded an effect of the enzymatic environment and specially, of the two Mg^{2+} ions resulting in better geometries likewise when the PM3 Hamiltonian was used. Therefore, this semiempirical method was thought to be the best to study the phosphoryl transfer reaction on the active site of the DHAK enzyme.

Most of results in this PhD Thesis have been obtained through the evaluation of free energy surfaces (see section 5.4) which requires the analysis of a large number of structures using QM/MM MD simulations. Therefore, calculations are normally limited to the employment of semiempirical methods being needed to reduce the errors related with their use. The way of doing it in this

work is based on the studies of Truhlar et al.,[179-181] and consist in applying an interpolated correction term to any value of the reaction coordinate, ξ, selected to produce the free energy surface (FES). Thus, a continuous new energy function is generated which corrects the potential on mean force (PMF) (see section 5.4):[182,183]

$$E = E_{LL/MM} + S[\Delta E_{LL}^{HL}(\xi)] \quad (5.18)$$

In equation (5.18), $E_{LL/MM}$ is the calculated energy where the QM subsystem is described with a low-level (*LL*) methodology, that is, with semiempirical methods, S means a spline under tension function[184,185] and ΔE_{LL}^{HL} is a correction term evaluated from the single-point energy difference between a high-level (*HL*) and a *LL* calculation of the QM subsystem. As pointed out above, the B3LYP method is selected for the *HL* energy calculation employing the 6-31G(d,p) basis set.

S is adjusted to a defined grid depending on the reaction step studied. Single-point energy calculations at *HL* are computed on optimized structures in the corresponding PESs performed at *LL*.

In the case of reaction steps analysed within 2D-PMFs, the correction term is expressed as a function of two coordinates, ξ_1 and ξ_2, resulting in this energy function:[186]

$$E = E_{LL/MM} + S[\Delta E_{LL}^{HL}(\xi_1, \xi_2)] \quad (5.19)$$

Where S is adjusted to an ensemble of points corresponding to the *HL* single-point energy calculations on geometries optimized with semiempirical methods, as previously explained.

5.2. Potential Energy Surfaces

The potential energy of a particular structure gives information about its stability. Differences between potential energies of different structures determine which one is more stable and thus, more likely to be detected experimentally.

A reactant system is described by the following Hamiltonian:[187]

$$\hat{H} = \hat{K}_e + \hat{K}_n + \hat{V}_{en} + \hat{V}_{ee} + \hat{V}_{nn} \quad (5.20)$$

Where:
$\hat{K}_e$ is the kinetic energy of the electrons,
$\hat{K}_n$ is the kinetic energy of the nuclei,
$\hat{V}_{en}$ is the potential energy of electrostatic interaction between electrons and nuclei,
$\hat{V}_{ee}$ is the energy of electrostatic repulsion between electrons,
$\hat{V}_{nn}$ is the energy of electrostatic repulsion between nuclei.

The potential energy surface (PES) arise in the context of the Born-Oppenheimer approximation[143] which allows the separation between the nucleus and the electron motion owing to the huge difference between their masses, as explained at the beginning of the section 5.1. Therefore, the electrons move much faster than the nuclei and thus, the latter is supposed to be static, while the electrons are moving at a specific positions around them.
If this approximation is applied to the time independent Schrödinger equation, the equation 5.1 was obtained where a solution for each nuclear configuration for the electronic system is achieved depending on the coordinates of the quantum and classical nuclei (R_α and R_M). Therefore, if the location of the nuclei is considered fix, $\hat{K}_n$ would be 0 and $\hat{V}_{nn}$ constant. Adding this internuclear repulsion energy ($\hat{V}_{nn}$) to the electronic energy $E(R_\alpha, R_M)$, the potential energy is obtained which is the total energy of the molecule or set of molecules corresponding to a fix configuration of the nuclei:[187]

$$V(R_\alpha, R_M) = E(R_\alpha, R_M) + V_{nn}(R_\alpha, R_M) \quad (5.21)$$

Then, a PES is a representation of the potential energy for different relative positions of the nuclei and each point represents a structure of the molecular system being possible going from reactants to products and thus, describing the reaction.[188]
For describing a given system using Cartesian coordinates, 3N variables (degrees of freedom) are required being N the total number of atoms of the system. The employment of Cartesian coordinates removes implicitly the rotational and translational normal modes, even so, when the energy is represented versus the intern coordinates, a PES of 3N-5 dimensions would be obtained which is impossible to visualize (3N-6 vibration modes or independent coordinates plus the energy). In a chemistry process exists an

intern coordinate or combination of these which govern the transition between the different species, that is, parameters that change in a substantial way during the reaction. It is possible to represent the potential energy according to this coordinate obtaining a very satisfactory approximation of the chemical reaction. This coordinate is called reaction coordinate which obviously depend on the nature of the chemical process that is being studied. In any case, it should be pointed out that some experience is needed to choose the best distinguished reaction coordinate in each case.
Once the distinguished reaction coordinate or coordinates are chosen, each point of the PES will be computed varying (but keeping frozen in each calculation) this or these coordinates and minimizing the rest of the coordinates. An example of a PES is shown in Figure 5.5.

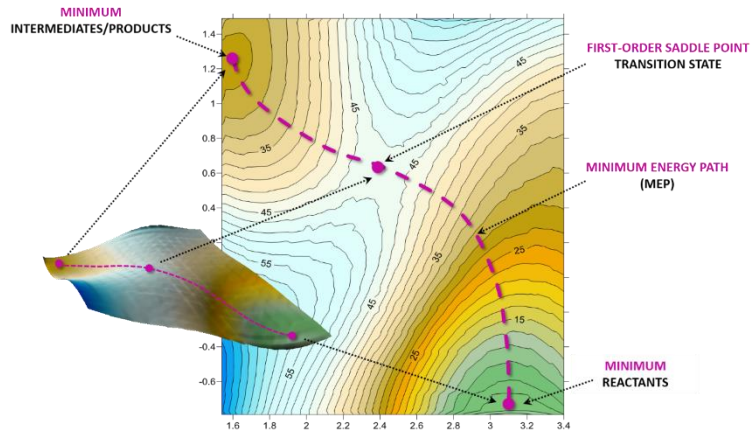


Figure 5.5. Potential energy surface (PES) in a 2-D and 3-D representation. The stationary points of interest have been pointed out as a dots and the minimum energy path from reactants towards products has been displayed as a dashed line.

As can be seen in Figure 5.5, the obtained PES allows to localize the different stationary points in the reaction process corresponding to reactants, products, intermediates and transition states.
In order to accomplish with the condition of stationary point, the gradient must be 0 and to characterize it in a PES, the Hessian should be known.
The eigenvalues of the Hessian can be determined by diagonalizing the matrix, that is;

$$|H - \lambda I| = 0 \qquad (5.22)$$

Where H is the Hessian, I is the identity matrix, λ are the eigenvalues and the straight lines denotes the determinant. There is an eigenvector associated with each eigenvalue satisfying the following equation:

$$F\vec{u}_i = \lambda\vec{u}_i \qquad (5.23)$$

Where F is the diagonalized Hessian matrix and $\vec{u}_i$ is the eigenvector which represents the principal curvature direction and the normal modes. There will be $3N$ eigenvalues and eigenvectors, being N the number of atoms in the system.

Depending on the results obtained after the diagonalization of the Hessian matrix, the most relevant stationary points obtained in a PES can be classified as minima or first-order saddle point (see Figure 5.5):

- Minima: All the eigenvalues obtained are positive. This means that an infinitesimal displacement of the geometry of the system along the direction defined by any of its eigenvectors, will lead to an increase in energy. The minima are the reactants, products and possible intermediates.

- First-order saddle points: All the eigenvectors obtained are positive minus one. If the displacement is along the eigenvectors associated to the positive eigenvalues, the energy will increase. However, an infinitesimal displacement along the eigenvector corresponding to the negative eigenvalue causes a decrease of the energy. This eigenvector associated to the negative eigenvalue is known as the transition vector and indicates the direction from the transition structure towards reactants and products.

Locating transition structures is more difficult than locating minima since it should be found a point which is a maximum in one direction but a minimum in all the others. Then the algorithm has to perform a discerning balance process between both types of search. The J. Baker algorithm, the one applied in the calculations of this PhD Thesis, which uses the first- and second-derivative information, is the most widely used and seems to be one of the most adequate methods available.

5.2.1. Intrinsic reaction coordinate (IRC)

A reaction path that connects reactants to products through the minimum energy can be followed in a given coordinate system as displayed in Figure 5.5. This reaction path is defined as the steepest descent path or minimum energy path (MEP) from the transition state down to the reactants and down to the products. If mass-weighted Cartesian coordinates are used, this path is called the intrinsic reaction coordinate (IRC).[189]

The IRC was proposed by Fukui in 1970[190,191] and represents the path that, departing from the transition structure, leads to the reactants and products, following the direction of the gradient.[159] In the initial point the transition structure has a zero gradient, thus the initial direction of the IRC should be previously specified by the transition vector (the unique negative eigenvalue of the hessian).

Mathematically the IRC can be obtained as the solution of the following differential equation system:

$$\frac{d\vec{x}}{ds} = \frac{-\vec{g}(\vec{x})}{|\vec{g}(\vec{x})|} \tag{5.24}$$

Where *s* is the path length from the transition structure or simply, the reaction coordinate and it is computed by:

$$(ds)^2 = \sum_i [(dx_i)^2 + (dy_i)^2 + (dz_i)^2] \tag{5.25}$$

Where the subscript *i* refers to each atom and *s* adopts positive or negative values depending on whether the displacement is towards products or reactants, respectively. Fukui[192,193] demonstrated that if mass-weighted Cartesian coordinates are used the requirement of following the gradient coincides with the path that would follow the system if, starting from the transition structure, it would go to reactants and products following a zero kinetic energy trajectory, that is, an infinitely slow trajectory:

$$q_i = x_i\sqrt{m_i} \tag{5.26}$$

Where x_i refers to an atom *i* Cartesian coordinate (x, y, z) and m_i to its mass. Therefore, the IRC is the reaction path which results from the solution of the

differential equation system (5.24) expressed in mass-weighted Cartesian coordinates. This is a useful tool which provides a certain way of knowing which reactants and products connect a transition state of a given reaction.

5.3. Molecular dynamics

The energy profile of a reaction in condensed media can be schematically represented as:

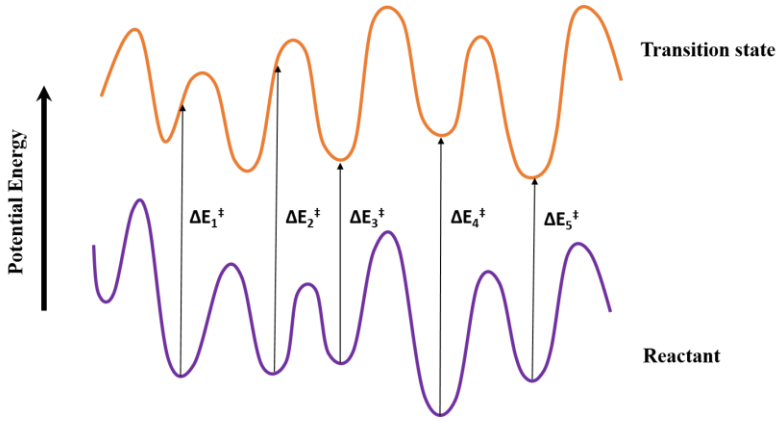


Figure 5.6. Schematic potential energy diagram of a reaction in a condensed media. Figure adapted from reference[10].

As it can be observed in Figure 5.6, in a reaction taking place in condensed media, there are many reactant structures, transition structures and product structures. This is due to the effect of the environment: there is a wide range of accessible structures which are minima or saddle points in the PES, with similar geometries and energies. Therefore, the properties of the transition state will be the average properties of all the transition structures, and the same occurs with the reactant and product states.

As shown in Figure 5.6, an energetic barrier ($\Delta E_1^{\ddagger}$) is obtained when one reactant structure and one transition state structure is picked up. This energy can be different if other structures for reactants and transition state are selected, ($\Delta E_2^{\ddagger}$), although they provide almost the same chemical information. Moreover, the same argument can be used for the reaction energies. Consequently, the isolated study of the PESs is not enough, and the evaluation of as much configurations as possible is required to obtain average properties,

which will be comparable to experimental data. Thus, it is necessary statistical simulations by means of MD or Monte Carlo simulations (MC) which allow the exploration of many different conformations existing in the system. MD simulations has been selected for the calculations in this PhD Thesis.
In order to carry out MD, the Newton's equation of motion for a particle should be known:

$$\vec{F}_i(t) = m_i\vec{a}_i = m_i\frac{\partial^2\vec{r_i}}{\partial t^2} \tag{5.27}$$

Once is known the force acting over a particular nucleus $\vec{F}_i(t)$, the position ($\vec{r}_i(t)$) and the velocity ($\vec{v}_i(t)$) in the previous instant, it is possible to calculate the new position after a time Δt. Among the different methodologies the *Verlet* method[194] has been selected. The standard *Verlet* method considers that if the positions of the atoms in the system at a time t are $\vec{r}_i(t)$ then the position of the atoms at time $t + \Delta t$ can be calculated from a Taylor expansion, leading to the following equations:

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \Delta t^2\frac{\vec{F}_i(t)}{m_i} \tag{5.28}$$

$$\vec{v}_i(t) \cong \frac{1}{2\Delta t}\left(\vec{r}_i(t + \Delta t) - \vec{r}_i(t - \Delta t)\right) \tag{5.29}$$

Equation 5.28 describes the positions of the particles at a time $t + \Delta t$ while equation 5.29 describes the velocities of the particles, $\vec{v}_i(t)$ at a time t. Nevertheless, the velocities at a time t can only be computed once the positions at a time $t + \Delta t$ have been obtained. Thus, at the beginning of the simulation, when t=0, a different formula is needed. To overcome this problem, the *Langevin Verlet* algorithm is employed,[194] which arises from a slight modification of the previous equations:

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t\vec{v}_i(t) + \frac{\Delta t^2}{2m_i}\vec{F}_i(t) \tag{5.30}$$

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{\Delta t}{2m_i}(\vec{F}_i(t) + \vec{F}_i(t + \Delta t)) \tag{5.31}$$

Within this algorithm, random forces depending on the velocities are introduced. In addition of having the necessary equations for knowing the position and the velocity of a particle at a particular time, it is necessary to fix some conditions and parameters required for conducting molecular dynamics simulations:

Conservation conditions

The accuracy of a simulation can be checked by the conservation conditions of certain properties, concretely the total momentum, $\vec{M}$, the angular momentum, $\vec{L}$ and the total energy;

$$\vec{M} = \sum_{i=1}^{N} \vec{p_i} = cnst. \tag{5.32}$$

$$\vec{L} = \sum_{i=1}^{N} r_i \wedge p_i = cnst. \tag{5.33}$$

Timestep

It should be considered which value is going to choose as the timestep. The factor which normally limits the upper size of the timestep is the nature of the highest frequency motions in the system which means that the timestep should be less than this, therefore, it will be possible to describe accurately the fastest vibration. In the practice, values of about 1 fs (1fs = 10^{-15}s) are found to be the largest reasonable when Verlet algorithms are employed.

Velocities initial values

Another point to consider is how to choose the initial values of the velocities for the atoms. A highly convenient procedure is to choose the velocities in such a manner that the system will have a particular temperature at the beginning of the simulation. From statistical thermodynamics it is known that the velocities of the atoms in a classical system are distributed according to the Maxwell-Boltzmann distribution. According to this, the probability of each component of the i_{th} atom velocity having a value between v and $v + dv$, if the temperature of the system is T is:

$$f(v)dv = \sqrt{\frac{m_i}{2\pi k_B T} exp\left(-\frac{m_i}{2k_B T} v^2\right) dv} \qquad (5.34)$$

Where k_B is the Boltzmann constant. The values of the velocities on the atoms can be selected by treating them as independent *Gaussian random variables* for the distribution defined in the equation 5.34 which has a mean value of zero and a standard deviation of $\sqrt{k_B T/m_i}$. Since the values are assigned randomly, the temperature of the system will not be exactly *T*, however, there is a popular expression which fix the temperature of the system by relating it with the average of the kinetic energy:

$$T = \frac{2\langle K\rangle}{N_{df} k_B} \qquad (5.35)$$

Where $\langle K\rangle$ is the averaged kinetic energy and N_{df} is the number of degrees of freedom of the system.

An instantaneous temperature, T_t, can be defined once the initial velocity values have been selected. The expression is the same that the previous one but removing the average:

$$T_t = \frac{2K}{N_{df} k_B} \qquad (5.36)$$

Thermodynamic ensemble

Working at constant volume (V), temperature (T) and total number of particles (N), is known as canonical or NVT ensemble. In order to keep constant the temperature an extern thermostatic bath is employed coupled with the system, as the originally proposed by Nosé and Hoover.[195,196] A coupling to an external system means that energy can be transferred into and out of the system. This transfer properly formulated allows the algorithm to control the temperature. Berendsen et al.[23] modify the equation of motion to accomplish the coupling, in the following way:

$$\vec{a}_i(t) = m_i \vec{F}_i(t) + \frac{1}{2\tau_i} \vec{v}_i(t) \left(\frac{T_{bath}}{T_t} - 1\right) \qquad (5.37)$$

Where T_{bath} is the temperature of the external bath, T_t is the instantaneous temperature defined in equation (5.36) and τ_i is the coupling constant which determines the strength of the coupling to the external bath.
This added term works as a frictional force, which makes that, when T_t is higher than the desired value, the force is negative, so the particles are slowed down and, therefore, the kinetic energy and temperature are reduced. Whereas, if T_t is lower than the required value, the invers process happens since the frictional force is positive and energy is supplied to the system.

5.4. Evaluation of Free Energies. Potential of Mean Force

Prediction of free energy barriers and reaction free energies has a huge importance for understanding enzymatic activity. In relation to the thermodynamics or the concepts of statistical mechanics, it can be affirmed that the free energy is the indispensable main driving force of enzymes in the thermodynamic limit. If the reaction free energy is negative, the forward reaction is more favourable while a positive value promotes the backward reaction.[11]
Working under the canonical ensemble conditions, the observable obtained for the free energy is the Helmholtz potential, A, which measures the "useful" work obtained from a system isolated from its surrounding environment at a constant temperature, and it is defined as:

$$A = -k_B T ln Q_{NVT} \tag{5.38}$$

Where Q_{NVT} is the partition function for a system formed by N indistinguishable particles for the working conditions mentioned. It gives the probability to find the system in a certain state:

$$Q_{NVT} = \frac{1}{h^{3N} N!} \int e^{-(\hat{H}(\vec{p},\vec{r})/k_B T)} dr dp \tag{5.39}$$

It is possible to calculate the free energy of a system from a simulation, rewriting the previous partition function as an average where Q_{NVT} is a property that depends on the energy of each configuration:

$$Q_{NVT} \propto \frac{1}{\langle e^{V/k_BT} \rangle} \tag{5.40}$$

It might be thought that a feasible way to calculate the free energy is to evaluate the average ($\langle e^{V/k_BT} \rangle$) along the trajectory from a molecular dynamics simulation, calculate the partition function and hence the free energy. However, this approach is impractical since the impossibility of exploring all the configurations allowing the system evolving freely. In this case it would be easier for the system to access to low-energy zones since simulations will preferentially explore the configurations with a lower potential energy because they have higher probabilities. To avoid this problem, the *Umbrella Sampling* technique,[197] proposed by J.P. Valleau and G.M. Torrie, is employed. This method allows to explore high-energy zones that would not be explored if the system was left to evolve freely, making sure that all regions along the reaction coordinate are sampled with similar probability.

Hence, on each MD simulation or window, an umbrella potential function is applied which restricts sampling to a limited range of values of the chosen coordinate, $\xi_{0,i}$, as shown in Figure 5.7a. This restriction is reached by placing a bias, an additional artificial energy term along the reaction coordinate.[11] A common choice is an harmonic potential form characterized by a force constant k_{umb} and a reference value of the coordinate whose value is changed at each window, $\xi_{0,i}^{ref}$.

$$V_{umb,i}(\xi_{0,i}) = \frac{1}{2} k(\xi_{0,i} - \xi_{0,i}^{ref})^2 \tag{5.41}$$

At this point, an observable associated to the free energy which depends on the reaction coordinate is needed. The Potential of Mean Force (PMF) is the thermodynamic concept associated to the evolution of a system along a coordinate and the free energy changes associated with this. The PMF, $U(\xi_0)$, has the same equation as the Helmholtz potential but the averages are made on all the degrees of freedom of the system except along the one that corresponds to the distinguished coordinate, ξ. It is defined as:

$$U(\xi_0) = c - k_B T ln(\int d\vec{r} \delta(\xi(\vec{r}) - \xi_0) e^{v(\vec{r})/k_BT}) \tag{5.42}$$

Where ξ_0 is the value of the degree of freedom ξ for which the PMF is computed, $\xi(\vec{r})$ is the function that relates ξ to the atomic coordinates $\vec{r}$, c is an arbitrary constant which includes all that is independent of ξ and Dirac delta function selects only those combinations of the atomic coordinates which give the reference coordinate value, ξ_0.
However, it is more common to write the PMF in terms of the ensemble average of the probability distribution function of the coordinate which is expressed as:

$$\langle \rho(\xi_o) \rangle = \frac{\int d\vec{r}\,\delta(\xi(\vec{r}) - \xi_0) e^{v(\vec{r})/k_B T}}{\int d\vec{r}\, e^{-v(\vec{r})/k_B T}} \tag{5.43}$$

then, the PMF is:

$$U(\xi_0) = c' - k_B T ln \langle \rho(\xi_0) \rangle \tag{5.44}$$

Where c' is another arbitrary constant.

To perform a correct computation of the whole density along ξ_0 (to obtain each value of $\langle \rho(\xi_0) \rangle$), a series of calculations should carry on for sampling in different regions of the space (or windows), but this regions should overlap in order to cover the whole configurational space and make the integration possible.

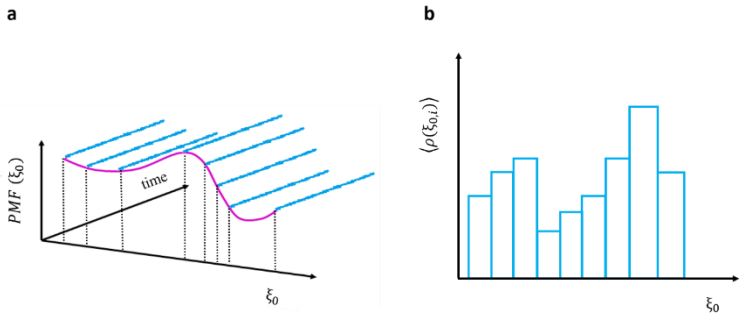


Figure 5.7. Representation of the Umbrella Sampling technique, where the energy profile is divided into several windows or MD simulations (a). Representation of the WHAM method where the complete distribution function is obtained from the distribution functions of each window (b).

Finally, the reconstruction of the full distribution function is performed from the separate distributions of each window, which is a fundamental step in the PMF calculation, obtaining the averaged free energy of the system, which, eventually, can be compared to the experimental data. The Weighted Histogram Analysis (WHAM),[198] elaborated by S. Kumar and co-workers, is an efficient procedure for doing this (see Figure 5.7b). It allows to construct an optimal estimate for the average distribution function from the biased distribution function for each window.

If it is supposed that there are N windows and each of one has an umbrella potential, $V_{umb,j}$, and an associated biased distribution function, $\langle\rho(\xi_{0,i})\rangle_{biased}$, the following equations will allow to calculate the average distribution function, $\langle\rho(\xi_0)\rangle$, and the PMF from a set of window distribution functions:[150]

$$\langle\rho(\xi_0)\rangle = \frac{\sum_{i=1}^{N} n_i \langle\rho(\xi_{0,i})\rangle_{biased}}{\sum_{j=1}^{N} n_j e^{-\frac{1}{K_B T}\left(V_{umb,j}(\xi_0) - PMF\left(\xi_{0,j}\right)\right)}} \tag{5.45}$$

$$e^{-\frac{1}{K_B T} PMF(\xi_{0,i})} = \int \langle\rho(\xi_0\rangle\, e^{-\frac{1}{K_B T} V_{umb,i}(\xi_0)} d\xi_0 \tag{5.46}$$

Where $n_{i/j}$ represents the number of employed data for calculating the probability distribution function of the corresponding window (frequency histograms in Figure 5.7b) and $PMF\left(\xi_{0,i/j}\right)$ is a window free energy. The coupling between the different windows is achieved by iterating in a self-consistent way. The procedure is to start by guessing a set of values for the free energies of each windows, $PMF\left(\xi_{0,j}\right)$ and with these, calculate $\langle\rho(\xi_0)\rangle$ for the complete range of ξ_0 from equation (5.45)[150]. This estimation of the distribution function is then used to determine the window free energies from equation (5.46) and the process is repeated until there are not changes either in the free energies or in the total density function.

5.5. Transition State Theory

The most important theoretical basis related with the enzymatic rate constant and therefore, capable of providing an estimation of the kinetic rate, is the Transition state theory (TST).[11] The TST was developed by Eyring, Evans and

Polany in 1935.[199,200] It is a semi-classical theory where the dynamics along the reaction coordinate are treated classically, while the perpendicular directions take into account the quantization of, for example, the vibrational energy.[137]
Since the generation of PES and the collection of a sufficiently great number of trajectories to realize a statistical average for the obtaining of the reaction rate is extremely difficult, a more straightforward theory is required. The TST theory allows the Born-Oppenheimer approximation and is based on the definition of a region in the space that contains all the nuclear configurations corresponding to the saddle point, dividing the PES of the reaction in two zones, reactants and products. Any structure placed in a very small distance from the saddle point region defined is called activated complex. The set of all the activated complex is denominated transition structure.
TST theory is based on the following postulates:[187,201]

- The reactant region of phase space is populated according to a Maxwell Boltzmann distribution corresponding to the temperature of the system. There is thermal equilibrium between reactants and TS during the entire reaction;

- The Born–Oppenheimer separation of electronic and nuclear motion is valid, thus, only one potential energy surface is involved. Processes in which electronic motion follows nuclear motion without an electronic state change are called electronically adiabatic, which strictly means that the electronic quantum numbers are conserved;

- The nuclear motion is classical;

- All trajectories crossing the transition state toward products must have originated on the reactant side and must cross the hypersurface only once. This condition is now usually called the ‘‘no recrossing’’ assumption. If the reaction coordinate were indeed globally separable, there would be no recrossing. Thus this condition is sometimes replaced by the assumption that the reaction coordinate is separable.

In this chemical reaction:

$$A \underset{}{\overset{K^{\ddagger}}{\rightleftharpoons}} A^{\ddagger} \xrightarrow{k} \text{Products} \qquad (5.47)$$

Where $A^{\ddagger}$ is the transition state of the first postulate, $K^{\ddagger}$ is the equilibrium constant between A and $A^{\ddagger}$, and *k* is the first-order rate constant. The reaction rate is:

$$v_{reaction} = v[A^{\ddagger}] = \frac{k_B T}{h}[A^{\ddagger}] \tag{5.48}$$

Bearing in mind the definition of the equilibrium constant, $K^{\ddagger}$, and its relationship with the Gibbs' free energy:

$$K^{\ddagger} = \frac{[A^{\ddagger}]}{[A]} \tag{5.49}$$

$$\Delta G^{\ddagger} = -RTlnK^{\ddagger} \tag{5.50}$$

the concentration of the species in the transition state is obtained:

$$[A^{\ddagger}] = [A]e^{(\frac{-\Delta G^{\ddagger}}{RT})} \tag{5.51}$$

and, comparing equation (5.48) to equation (5.51):

$$v_{reaction} = \frac{k_B T}{h}[A]e^{(-\frac{\Delta G^{\ddagger}}{RT})} = k^{TST}[A] \tag{5.52}$$

The rate constant for the reaction will be:

$$k^{TST}(T) = \frac{k_B T}{h} e^{(-\frac{\Delta G^{\ddagger}}{RT})} \tag{5.53}$$

k^{TST} from equation (5.53) provides the crossover velocity in only one sense through the division surface of the TS. According to this expression, the reaction rate can be estimated if it is known the free energy between reactants and the transition state, $\Delta G^{\ddagger}$. It should be emphasised that different choices of the reaction coordinate will result in different realisations of TST and different activation free energies. TST is designed to provide the reaction rate constant, and it does not describe the way from reactants to products.[11]

5.5.1. Variational Transition State Theory (VTST)

The approximations pointed out in the previous postulates imply that TST is not able to explain different phenomena which are mainly two: tunnelling and recrossing:[187,201,202]

Tunnelling: As stated above, the treatment of the system motion along the reaction coordinate is classical. It would be more accurate to describe this motion quantumly which consider the possibility of tunnelling.[203] This phenomenon allows the reaction to occur even when there is not energy enough to overcome the energetic barrier (see Figure 5.8a). This is the case of reactions where very light atoms are involved, such as hydrogen atoms. Therefore, there are trajectories that according to TST have no possibility of reaching products from reactants while from the quantum point of view, they have non-zero probability of evolving into products.
Recrossing: This limitation arises from the fact of having assume that all the trajectories that overcome the dividing surface from reactants towards products are reactive. Nevertheless, reaction dynamic studies prove that once the dividing surface is reached, the trajectory could go back again to reactants producing a recrossing phenomenon, as shown in Figure 5.8b. Depending on the reaction system, recrossing can be more or less relevant.

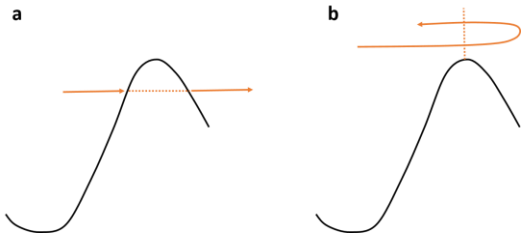


Figure 5.8. a; Tunnel effects consider a certain probability for some particles to react even when their energy is less than the barrier. b; The system, after crossing the dividing surface towards products, could get back to reactants.

Nevertheless, there is a way of quantifying these effects though the variational transition state theory (VTST) which relates the actual rate for the reaction to that obtained from TST:[202,204-207]

$$k(T) = \Gamma(T,\xi)k^{TST}(T) \qquad (5.54)$$

Being $k(T)$ the rate constant of a reaction and $\Gamma(T,\xi)$ the generalized transmission coefficient which contains the product of the two contributions at a temperature *T*:

$$\Gamma(T,\xi) = \kappa(T)\gamma(T,\xi) \tag{5.55}$$

Where $\kappa(T)$ emerges from the contribution of quantum mechanical tunnelling and $\gamma(T,\xi)$ is the coefficient responsible of the rate correction from recrossing of reactive trajectories.

Since some years from nowadays, there is a debate about whether the dynamical effects have a relevant role in catalysis.[11,208-212] It is easy to think that some dynamical effects can affect for example, the motion of reactant fragments to be closer and properly oriented in order to react and can also origin conformational changes resulting in an adequate electrostatic distribution to the reaction take place. Nevertheless, in 2010, Warshel and Kamerlin[208] and six years later, in 2016, Warshel and Bora[213] have emphasized in the need of a clear and specific definition of dynamical effects since some authors, supporters of the importance of dynamical effects in catalysis, have not been clear about the meaning of this term, leading to confusion. In these reviews authors proposed some definitions, starting from the point that as atoms are always moving in any chemical process, it does not have sense to classify the movement of the atoms as "dynamical effects". For example, they affirm that a dynamical effect is produced when the probability of reaching the TS does not follow the Boltzmann distribution, therefore, there will be non-Boltzmann motions that can be assorted as real dynamical effects. Thus, these effects require deviations from the original TST being the transmission factor significantly smaller than one. Regarding to the existent debate previously mentioned, in some works dynamical effects are mentioned to contribute in enzymatic catalysis in a significant way.[134,214,215] Among the reasons given by the authors are that the long MD simulations provide the suitable conformation to produce the catalytic reaction or that the trajectories are capable of guiding the reaction toward a concerted or a stepwise mechanism. On the other hand, other authors point out that dynamics doesn't affect in a relevant way, the rate constant of the enzymatic reaction.[208-211,216] These authors explain that using a clear and physically-based definition of the dynamical proposal, the facts observed by the supporters of dynamical effects can properly explain without appealing to this term. Therefore, in these cases authors support the

employment of the TST theory as sufficiently adequate to describe rate constants. It should be also mentioned that, normally, the transmission factor has a value of 0.8 in enzyme and 0.6 in solution[11], which means that they are too similar for considering that these effects can have a relevant impact in catalysis.

Specially in the last years, the contribution of quantum tunnelling has been accepted to occur in almost all enzymes that involve hydrogen transfer, making the reaction dependent on the width of the barrier as well as on the height, according to some authors.[217,218] Reactions involving quantum tunnelling has been also related with dynamical effects, specifically as regards to environmental motions, causing and impact in the rate constant.[211] There are authors who suggest that the presence of tunnelling prevent from the use of TST in enzymatic reactions.[219,220] Nevertheless, the phenomena is not new and is common in chemical reactions in solution.[221] In addition, there are studies that affirm that the quantum mechanical contributions, are similar for the reaction in solution and in the enzyme and therefore, they are not decisive in enzymatic catalysis.[5,33,216]

All in all, either the tunnelling or the recrossing trajectories can have a minor effect on reactions involving the transfer of heavy particles as the phosphoryl group studied in the present PhD Thesis. Consequently, these effects have not been estimated in the reported studies

6. Conclusions and outlook

In the present PhD Thesis, the molecular mechanism of DHAKs has been tried to be elucidated. First of all, in order to determine the catalytic power of these enzymes, a study of the analogous uncatalysed reaction in solution was considered. This first study was also used to test different methods in order to optimize the computational protocol to be applied in the following studies. Therefore, the phosphoryl transfer from ATP to Dha in aqueous solution employing QM/MM methods has been analysed using different Hamiltonians for describing the QM region. A concerted, stepwise and assisted by the solvent mechanisms were studied, although the latter presented higher potential energies than the others. Therefore, it was not considered for further calculations. FESs for the two most favorable mechanisms revealed a significant dependency on the results with the QM method employed with relevant consequences. The first conclusion was that AM1d/MM and PM6/MM levels were able to describe the concerted mechanism but any effort to describe it at PM3/MM level was in vain. In the case of the stepwise mechanism, AM1d/MM and PM3/MM described the two stages while the second step was not possible to be located with PM6/MM in terms of free energies being the intermediate (where the Dha-P is protonated) more stable than the final products. Regarding the comparison of the FESs topology, in the case of the concerted mechanism, the TS was approximately located in the same quadratic region at AM1d/MM and PM6/MM levels resulting in similar degrees of phosphoryl and proton transfer. However, in the first step of the stepwise mechanism, the protonation in one of the phosphate group oxygen of ATP took place firstly, followed by the phosphoryl transfer in the case of AM1d/MM while at PM6/MM level the opposite process occurred, in both cases within a TS. Using PM3/MM, this first stage of the stepwise mechanism was described following a similar reaction path than in the AM1d/MM calculations but within two steps clearly differentiated, with the presence of an intermediate. In this intermediate the ATP was protonated in its terminal phosphate group. The TS structures obtained in the second stage of the stepwise mechanism using these latter Hamiltonians were also similar although there were divergences in the energetics. When *spline* B3LYP/MM corrections were carried out on the FESs, significant changes were produced in the topology of the surfaces for the PM3/MM and PM6/MM calculations. When the high level corrections were applied to the concerted PM6/MM mechanism the transfer of the proton took place before de phosphoryl transfer, unlike in the original surface. AM1d/MM free energy surfaces corrected at high level of theory resulted in a similar topologies than the original ones. The

representation of the entire free energy profiles corresponding to the explored mechanisms using the different Hamiltonians at semiempirical/MM and semiempirical:B3LYP/MM levels showed lower energy barriers for the PM3/MM stepwise mechanism.

Since the *spline* corrections seems to provide different results depending on the original semiempirical/MM free energy surface, PES at B3LYP/MM level were conducted in order to look for the most reliable path. The obtained PES resembled more to the AM1d/MM results justifying also the small changes observed when these original free energy surfaces were corrected. Moreover, these B3LYP/MM PESs showed not very stable products in both concerted and stepwise mechanisms. These results could explain the fact that the concerted mechanism and the second step of the stepwise mechanism were not found through PM3/MM and PM6/MM levels respectively.

As the next step of this PhD Thesis, the reaction was investigated in the active site of the enzyme. Thus, a detailed theoretical study of the phosphoryl transfer mechanism was carried out in the DHAK from *E. coli*. Mainly, two phosphorylation paths have been explored: a *substrate-assisted mechanism*, where the activation of the Dha for the nucleophilic attack is due to the ATP substrate, and an *asp-assisted mechanism* where this activation is produced by the conserved Asp109A residue. Obtained results suggested that the former is kinetically more favorable than the later where the Asp109A was proved to play a structural role placing the Dha and other residues in the proper orientation for the reaction to take place. This conclusion was reinforced by the energetics of the intermediates obtained after the phosphate transfer where the one corresponding to the *substrate-assisted mechanism* was more stable than the equivalent one located in the *asp-assisted mechanism*. Interesting conclusions were derived from the evaluation of FESs. The key step of the molecular reaction corresponding to the proton and the phosphoryl transfer was shown to be stepwise at PM3/MM level, within a presence of an intermediate, and concerted at B3LYP/MM level in both explored mechanisms. Nevertheless, the analysis of the representative snapshots of located TS structures obtained at PM3/MM and B3LYP/MM were structurally similar in each one of the mechanisms regarding to the phosphoryl transfer process; while in the *substrate-assisted* path the TS was characterized by being in an earlier stage of the procedure, in the *asp-assisted* path was in an advance stage of the reaction. The analysis of the role of the surrounding residues showed that there were some of them responsible of stabilizing the TS within both methods, in particular, Gly79,B Thr79A, Ser80A, Arg179B, one of the

Mg^{2+} cations. Asp109A also played a structural role in the case of the *substrate-assisted mechanism* properly orientating the Dha and some residues of the active site.

It should be pointed out some effects regarding the differences in the energetics depending on the employed method. In the case of the first step of the reaction mechanism, corresponding to Dha binding to the enzyme and the proton transfer from His56A to Dha, *spline* corrections at B3LYP:PM3/MM rendered different reaction paths than the original PM3/MM surface. While in the original surface a stepwise mechanism was displayed, two competitive asynchronous processes were observed in the corrected surface. Corrections in the phosphoryl transfer step of the *substrate-assisted mechanism* likewise revealed some changes. According to the B3LYP:PM3/MM surface, the phosphoryl transfer takes place firstly, followed by the proton transfer to the oxygen of ATP. This is just the opposite timing obtained in the original PM3/MM surface. Evaluation of the PM3/MM free energy barriers in the *substrate-assisted mechanism* determined the stage corresponding to the release of the phosphorylated substrate as the rate limiting step, with a barrier of 21.6 $kcal \cdot mol^{-1}$. In contrast, the rate limiting step in the *substrate-assisted mechanism* derived from B3LYP/MM calculations was the phosphoryl transfer being the barrier 30.0 $kcal \cdot mol^{-1}$. However, it must be kept in mind that a change in the value of the barrier could be expected if statistical simulations were carried out to generate FESs.

In collaboration with the group of Prof. Eduardo García-Junceda of the Institute of General Organic Chemistry of CSIC (Madrid, Spain), an attempt of tuning DHAK enzyme to use poly-P as the phosphoryl donor has been carried out through protein engineering techniques in combination with computational methods. Theoretical calculations based on MD simulations and QM/MM calculations allowed to study the interaction energies in the wild type and the mutant form of the kinase, to analyse the effects caused by the mutation, Glu526Lys, on the binding on the poly-P. The average values of the computed interaction energies showed a higher binding energy between the poly-P and the enzyme in the mutated system. Analysis of the contribution of each single residue and the Mg^{2+} ions in the interaction energy (electrostatic plus van der Waals) with the poly-P proved that this energy increased in the case of the residues Lys382A and Arg475A and in a remarkable way, with the Mg^{2+} ions. Therefore, it was concluded that the mutation did not provoke an interaction attractive effect between the poly-P and the new mutated residue Lys, but rather an indirect consequence, pushing the poly-P towards de active

site and favoring the reaction to take place. This approaching to the active site also implied an enhancement of the repulsive energy of the Asp residues coordinating with the Mg^{2+} ions. Results from the Root Mean Square Fluctuation (RMSF) evaluated from the MD simulations showed that the mutated system presented a higher mobility, in particular, in the flexible *loop*, where the mutation was located. Specifically, the mobility of Arg519, Ala520 and Ser521 was detected to be higher after the mutation. This fact could be related with a better accommodation of the poly-P in the active site and it suggested that attempts to modify the specificity to poly-P should be focused in these residues.

The final study of the present PhD Thesis was the chemical step of the phosphoryl transfer reaction from the poly-P to the Dha in DHAK from *C. freundii* in the wild type and in the most active mutant. The *substrate-assisted mechanism* and the *asp-assisted mechanism* were analysed. The exploration of the PM3/MM potential energy surfaces showed the *substrate-assisted mechanism* as the most favorable reaction path in both, the wild type and the mutated protein. Location of the key stationary points at B3LYP/6-31G(d,p) level of theory in the most favorable mechanism displayed similar potential energy barriers for the two systems showing that the mutation has tiny effects on the chemical step, thus confirming our preliminary conclusions deduced from our previous study.

The results presented in this PhD Thesis provide valuable information on one of the most important reactions in Nature. The conclusions can be used not only to have a better understanding of the reaction, but also for further studies in other related kinases. The specific details of every system may be taken advantage to develop more potent drugs with less side-effects. On the other side, continuation of the tuning of DHAK can be based on the proposal and exploration of additional mutations that favor the binding step of the poly-P and reduce the activation free energy barrier of the chemical step. Successful results could open the door to the use of cheaper and more environmental friendly catalysts in industry, or the optimization of protocols to design new high value chemicals.

7. Publications

7.1. Study of the phosphoryl transfer reaction from ATP to Dha in aqueous solution

A computational study of the phosphoryl transfer reaction between ATP and Dha in aqueous solution.

Isabel Bordes, José Javier Ruiz-Pernía, Raquel Castillo and Vicent Moliner

Published in *Organic & Biomolecular Chemistry* **2015**, 13, 10179-10190

Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

CrossMark click for updates

Cite this: *Org. Biomol. Chem.*, 2015, **13**, 10179

A computational study of the phosphoryl transfer reaction between ATP and Dha in aqueous solution†

I. Bordes, J. J. Ruiz-Pernía, R. Castillo* and V. Moliner*

Received 28th May 2015,
Accepted 12th August 2015

DOI: 10.1039/c5ob01079a

www.rsc.org/obc

Phosphoryl transfer reactions are ubiquitous in biology, being involved in processes ranging from energy and signal transduction to the replication genetic material. Dihydroxyacetone phosphate (Dha-P), an intermediate of the synthesis of pyruvate and a very important building block in nature, can be generated by converting free dihydroxyacetone (Dha) through the action of the dihydroxyacetone kinase enzyme. In this paper the reference uncatalyzed reaction in solution has been studied in order to define the foundations of the chemical reaction and to determine the most adequate computational method to describe this electronically complex reaction. In particular, the phosphorylation reaction mechanism between adenosine triphosphate (ATP) and Dha in aqueous solution has been studied by means of quantum mechanics/molecular mechanics (QM/MM) Molecular Dynamics (MD) simulations with the QM subset of atoms described with semi-empirical and DFT methods. The results appear to be strongly dependent on the level of calculation, which will have to be taken into account for future studies of the reaction catalyzed by enzymes. In particular, PM3/MM renders lower free energy barriers and a less endergonic process than AM1d/MM and PM6/MM methods. Nevertheless, the concerted pathway was not located with the former combination of potentials.

Introduction

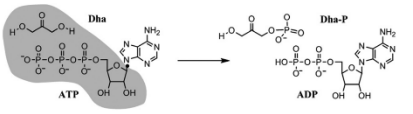
Phosphoryl transfer, a chemical process that consists of the transfer of the phosphoryl group from a phosphate ester or anhydride to a nucleophile,[1] is involved in a wide range of biological processes, from energy and signal transduction to the replication of genetic material.[2,3] Protein kinases catalyze the transfer of phosphate groups from adenosine triphosphate (ATP) to different substrates. ATP functions as a chemical energy carrier[4] since it is a derivative of the nucleic acid adenine that has three high-energy phosphate bonds at the 5′ position. It is the most important biological phosphoryl donor required for many enzymatic reactions.[5]

When a phosphate group is transferred from ATP to one of the simplest carbohydrates, for example dihydroxyacetone (Dha), Dha-P is obtained (see Scheme 1).[6] Dha-P is an intermediate for the synthesis of pyruvate[7] and it is a very important C_3 building block in nature since it is used as a phosphoryl donor in several enzyme-catalyzed aldol reactions.[8] The phosphorylation of Dha is catalyzed in nature by the action of the dihydroxyacetone kinase enzyme.

The function of kinases has been observed to be deregulated in many human pathological disorders such as cancer. Accordingly, a large number of kinase inhibitors have been developed.[9,10] Nevertheless, despite their multiplicity only a part of the kinome has been so far targeted with some specificity and potency, leaving open a broad range of options for

Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castellón, Spain. E-mail: rcastill@uji.es, moliner@uji.es

† Electronic supplementary information (ESI) available: PES of the concerted and stepwise mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at AM1d/MM (Fig. S1), PM3/MM (Fig. S2) and PM6/MM (Fig. S3). The AM1d/MM, PM3/MM solvent assisted mechanism (Fig. S4); structures of stationary points of the solvent assisted mechanism at the AM1d/MM level (Fig. S5) and structures of stationary points of the solvent assisted mechanism at the PM3/MM level (Fig. S6). Potential energies of the mechanisms at AM1d/MM, PM3/MM and PM6/MM (Table S1) and corrected energies at B3LYP (Table S2). Cartesian coordinates of QM atoms of located stationary points. See DOI: 10.1039/c5ob01079a



Scheme 1 Schematic reaction between ATP and Dha in aqueous solution, generating ADP and Dha-P. Grey region contains atoms treated quantum mechanically while the adenosine group and water molecules (not shown for clarity purposes) are represented by classical force fields. Link atom in ATP is represented as a dot.

future research. Moreover, after decades of research, the molecular mechanisms of phosphoryl transfer reactions remain a hot topic of debate in the literature,[11] including the controversy of whether the non-enzymatic and the enzymatic reaction proceed by the same molecular mechanism. Remarkably, these reactions present very small non-enzymatic rates and thus require enormous rate accelerations from biological catalysts.[11] The large changes in the charge distribution between reactants and transition states (TSs) offer a good advantage for the reaction to be catalyzed[12] and, consequently, a deep knowledge of the reaction in solution becomes even more relevant. Finally, from the computational point of view, a study of the reaction in solution, apart from its interest to quantify the catalytic proficiency of the enzyme, can be used for the calibration and validation of the theoretical approaches to be used in subsequent studies of the corresponding enzyme-catalyzed reaction.[3] In this regard, and as remarked by Kamerlin, an additional difficulty for the theoretical study of the phosphoryl transfer reactions is due to the availability of low-lying d-orbitals on the phosphorus atom which allows the existence of phosphorus pentavalent species as intermediates. Moreover, phosphoryl transfer reactions can potentially occur through multiple equally viable mechanisms, even in solution.[2,3]

The phosphoryl transfer from phosphate derivatives to hydroxide ions or to water, namely the hydrolysis reaction, has received a lot of attention in the past few years.[2,3,13–16] Regarding the general mechanisms of phosphate derivative hydrolysis, a debate has been established around whether the reaction proceeds by an associative or dissociative mechanism from the early systematic work of Florian and Warshel.[17] As stated by Warshel and co-workers, some recent computational studies failed to consider the clear distinction between these two paths.[17] Another question of debate is whether the proton from the attacking water molecule is transferred to the terminal phosphate oxygen in a direct way or assisted by additional water molecules.[18] This latter possibility was originally explored by Hu and Brinck.[19] More recently, the competition between dissociative/solvent-assisted and associative/substrate-assisted pathways for a number of phosphate and monoester hydrolysis has been examined by Williams, Kamerlin and co-workers.[20] The authors also considered both direct proton transfer and proton transfer *via* an intervening leaving group, showing the difference as negligible, and that the substrate-assisted mechanism is only favored with poor leaving groups. Nevertheless, the specific reaction between ATP and Dha has not been studied yet by means of computational methods. The study of this reaction in solution can provide the bedrock for the understanding of the basic chemistry of Dha kinases.

Density functional theory (DFT) and *ab initio* methods have been used in many computational studies of phosphoryl transfer reactions in the gas phase or with implicit solvent models.[17,21–23] Nonetheless, the remarkable computation cost of these methods prevents their use to perform MD simulations, which have been shown to be an effective approach to get reaction free energy profiles. An alternative is employing semiempirical methods to treat the QM sub-set of atoms in hybrid QM/MM computational schemes. In the last few years, a modified semiempirical Hamiltonian AM1/d-PhoT (hereafter simply named as AM1d) has been developed to model phosphoryl transfer reactions.[24] AM1d incorporates d-extension for the phosphorus atom and modified AM1 parameters for oxygen and hydrogen atoms, while the remaining atoms are described at the AM1 level. This semiempirical potential has been shown to provide results in very good agreement with high-level DFT calculations in the study of the hydrolysis of phosphorus compounds.[24 27]

Another semiempirical method that is typically 3–4 orders of magnitude faster than DFT methods, the PM3 method,[28] has shown to produce a huge stabilization in the energy of phosphorane compounds, resulting from the use of a minimal valence basis.[24] Nevertheless, and despite its accuracy to model phosphoryl transfer reactions is limited because of the lack of d orbitals,[16,24] due to its low computational cost, it can be used in order to perform statistical simulations to rapidly explore free-energy surfaces and to establish in a reliable yet computationally feasible way an optimal computational protocol.[16] Several modifications that have been made to the NDDO core–core interaction term have resulted in a more complete parameter optimization called PM6.[29] The most important of these changes were the addition of d orbitals to main-group elements and the introduction of diatomic parameters.[30]

A proper comparison between different semiempirical Hamiltonians combined with a common MM force field would allow determining which is the most appropriate one to be used when running long QM/MM MD simulations. Within this target in mind, a theoretical study of the phosphoryl transfer reaction from ATP to Dha in aqueous solution, carried out by means of hybrid QM/MM potentials, is reported in the present paper. Different semiempirical Hamiltonians such as AM1d, PM3 and PM6 have been employed to describe the QM region of the system, while the standard B3LYP hybrid functional has been considered as the reference QM method. As stated above, this study will be the basis for further QM/MM studies of the reaction in the active site of Dha kinases.

Computational methods

Building the system

The initial coordinates of ATP and Dha were taken from the X-ray structure of *Escherichia coli* dihydroxyacetone kinase, with PDB entry 3PNL.[7] The original ADP molecule was modified to ATP by adding a phosphate group. Then, the two species, ATP and Dha, were solvated with a pre-equilibrated cubic box of water molecules of side 55.8 Å. Any water molecules with an oxygen atom lying within 2.8 Å of any non-hydrogen atom of the two solute species were removed. For all simulations, atoms belonging to molecules found at a distance less than or equal to 25 Å of one of the phosphorus atom of ATP were defined as flexible. The rest of the atoms were kept frozen.

In order to run the simulations with hybrid QM/MM potentials, the sub-set of atoms treated quantum mechanically are

those of Dha and part of ATP (phosphate groups and ribose ring, as shown in Scheme 1). When exploring the possible mechanisms assisted by solvent, a water molecule was also treated quantum mechanically. To saturate the valence of the QM/MM frontier we used the link atom procedure,[31] placing this atom between C of the ribose molecule and N of the adenine molecule (see Scheme 1). The number of QM atoms is 41, while the system with water molecules contains a total of 17 425 atoms.

The part of the ATP not included in the QM region was described by means of the OPLS-AA[32] force field and water molecules with the TIP3P force field,[33] as implemented in the fDYNAMO library.[34] AM1d/MM and PM3/MM calculations have been carried out with fDYNAMO whereas PM6/MM and B3LYP/MM calculations were performed by combining fDYNAMO with the Gaussian09 program.[35] In the case of the DFT/MM calculations, the employed basis set was 6-31G(d,p).

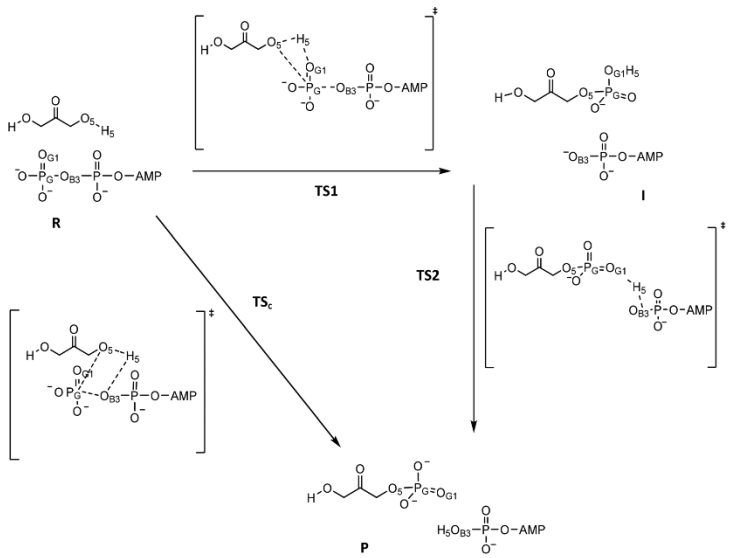
Simulations

First of all, the system was optimized using a combination of the steepest descent method and lbfgsb algorithm after relaxation by means of 100 ps of Langevin Dynamics using the NVT ensemble at 300 K with a time step of 1 fs. A switched cutoff, from 14 to 16 Å, was employed for all non-bonded interactions between MM atoms (including the frozen atoms), while the QM region was allowed to interact with every flexible MM atom.

Two phosphorylation reaction paths have been explored without direct participation of water molecules (see Scheme 2); a concerted mechanism and a stepwise mechanism.

The former is the direct transfer of a phosphate group from ATP to Dha, simultaneously with a proton transfer from Dha to the resulting ADP. In the latter the transfers are considered in two steps.

To explore both reaction mechanisms, potential energy surfaces (PESs) were computed by scanning the appropriate combination of the interatomic distances that define both transfers. Later free energy surfaces were computed in terms of potentials of mean force (PMFs).[36] In particular, a two-dimensional QM/MM PMF (2D-PMF) was computed to study the concerted mechanism using, as distinguished reaction coordinates, the distance between the phosphor PG atom of ATP and the O5 atom of Dha, d(PG–O5), and the antisymmetric combination of the distances describing the hydrogen transfer from Dha to ATP: d(O5–H5)–d(H5–OB3). In the case of the stepwise mechanism, two different PMFs have been explored. In the first one, a 2D-PMF was generated using the distance d(PG–O5) as one of the reaction coordinates and the antisymmetric combination d(O5–H5)–d(H5–OG1) as the other one. The second PMF was traced using the antisymmetric



Scheme 2 Representation of the two possible reaction mechanisms corresponding to the phosphoryl transfer from ATP to Dha.

combination *d*(OG1–H5)–(H5–OB3) as the distinguished reaction coordinate, obtaining a 1D-PMF. The umbrella sampling approach[37] was used to constrain the system along the selected values of the reaction coordinates by employing a force constant of 2500 kJ mol^{-1} Å^{-2}.

The probability distributions were put together by means of the weighted histogram analysis method (WHAM)[38] to obtain the full probability distribution along the reaction coordinate.

The values of the force constant used for the harmonic umbrella sampling were determined to allow full overlap of the different windows traced in the PMF evaluation, but without losing control over the selected coordinate. Each window consisted of 15 ps of equilibration followed by 20 ps of production. The Verlet algorithm was used to update the velocities. The initial structures in each window of the 2D-PMFs were selected from the corresponding previously generated PESs at the corresponding level of theory, whereas the 1D-PMFs were performed starting from the transition state structure.

Because of the large number of structures that must be evaluated during free energy calculations, QM/MM calculations are usually restricted to the use of semiempirical Hamiltonians.

In order to reduce the errors associated with the quantum level of theory employed in our simulations, following the studies of Truhlar *et al.*[39–41] a spline under tension[42,43] is used to interpolate this correction term at any value of the reaction coordinates, ξ_1 and ξ_2 in the case of the 2D-PMFs, or only one coordinate, ξ, for the 1D PMF. Thus, for correction of the 2D-PMFs, we obtain a continuous function in a new energy function to obtain corrected PMFs:[44,45]

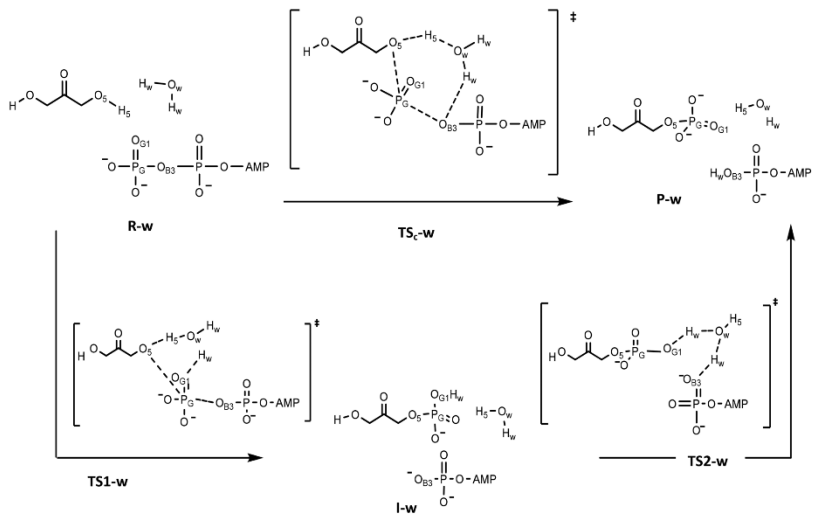
$$E = E_{\mathrm{LL/MM}} + S[\Delta E^{\mathrm{HL}}_{\mathrm{LL}}(\xi_1\xi_2)] \quad (1)$$

where S denotes a two-dimensional spline function, and its argument is a correction term evaluated from the single-point energy difference between a high-level (HL) and a low-level (LL) calculation of the QM subsystem. The selected semiempirical Hamiltonians were used as the LL method, while the B3LYP/6-31G(d,p) method was selected for the HL energy calculation. S is adjusted to a defined grid depending on the reaction mechanism studied and the semiempirical Hamiltonian employed. HL single energy calculations are computed on optimized geometries obtained in the corresponding PESs at LL. In the case of the 1D-PMFs, the second step of the stepwise mechanism, the correction scheme is defined as:

$$E = E_{\mathrm{LL/MM}} + S[\Delta E^{\mathrm{HL}}_{\mathrm{LL}}(\xi)] \quad (2)$$

and S is adjusted to a set of points corresponding to the HL single energy calculations on geometries optimized at AM1d and PM3 methods, as previously explained. HL calculations were carried out using the Gaussian09 program.

The possible solvent-assisted mechanisms have been explored in terms of PESs since, as demonstrated in the next section, the obtained energy barriers were much higher than



Scheme 3 Representation of the two possible solvent assisted reaction mechanisms corresponding to the phosphoryl transfer from ATP to Dha.

the ones obtained without the participation of explicit water molecules. Then, the much more expensive computing of free energy surfaces was considered as not required to discard these possible paths. Due to the large amount of key coordinates involved in these mechanisms, several anti-symmetrical combinations of distances were tested. The most efficient way to control the process was achieved using, as distinguished reaction coordinates, two antisymmetric combination of distances: the distances describing the hydrogen transfer from Dha to the quantum water molecule, *d*(O5–H5)–*d*(H5–Ow), and the antisymmetric combination describing the transfer of the phosphate group, *d*(OB3–PG)–*d*(PG–O5) (see Scheme 3).

Results and discussion

As stated in the previous section, the first step in our study was the exploration of the PESs corresponding to the concerted and stepwise mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution. A QM/MM scheme was used, with the QM region treated with different semiempirical Hamiltonians: AM1d, PM3 and PM6. Then, the resulting PESs (see Fig. S1–S3 of the ESI†) were used to generate the free energy surfaces in terms of 1D- and 2D-PMFs. The results are shown in Fig. 1 (AM1d/MM), Fig. 2 (PM3/MM) and Fig. 3 (PM6/MM), while averaged values of selected interatomic distances of the states located along the reaction paths (reactants, TSs and product states) are listed in Table 1. As observed in Fig. S4,† the PESs obtained after exploring the solvent-assisted mechanism render potential energy barriers much higher than the ones obtained without the participation of explicit water molecules and, consequently, the 2D PMFs were considered as not necessary to be computed for this mechanism (geometries of stationary points located on the PES of this mechanism are reported in Fig. S5–S7 of the ESI†).

Concerted mechanism

Analysis of Fig. 1–3 reveals that different Hamiltonians can render not only different values of activation and reaction energies, but also significant different reaction mechanisms. The first and most dramatic difference when exploring the concerted mechanism is that any attempt to locate a reaction path with the PM3/MM method was unsuccessful. On the contrary, the concerted TS obtained by means of AM1d/MM and PM6/MM levels are quite similar, according to the averaged values of the distances PG–O5, PG–OB3, O5–H5 and H5–OB3 listed in Table 1. As observed, both methods describe a TS where the breaking PG–OB3 bond and the PG–O5 forming bond are in a very advanced stage of the reaction, while the H5 transfer would present the opposite behavior being in an earlier stage of the process.

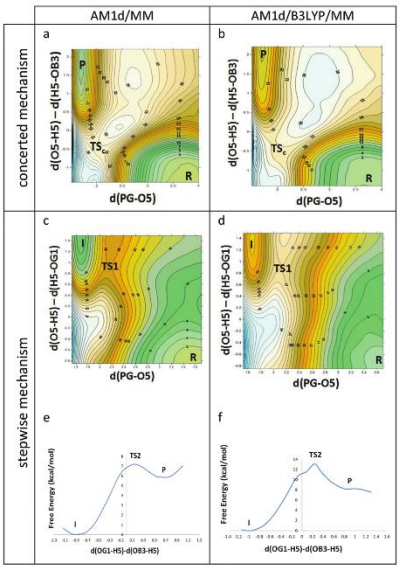


Fig. 1 Free energy surfaces of the concerted and stepwise mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution, obtained as 1D- and 2D-PMFs at AM1d/MM level (panels a, c and e) and with spline corrections at the B3LYP/MM level (panels b, d and f). Energies are given in kcal mol^{-1} and distances in Å.

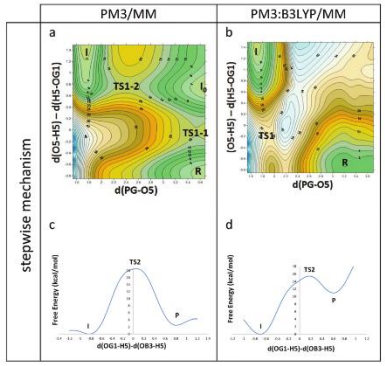


Fig. 2 Free energy surfaces of the stepwise mechanism of the phosphoryl transfer reaction from ATP to Dha in aqueous solution, obtained as 1D- and 2D-PMFs at the PM3/MM level (panels a and c) and with spline corrections at B3LYP/MM level (panels b and d). Energies are given in kcal mol^{-1} and distances in Å.

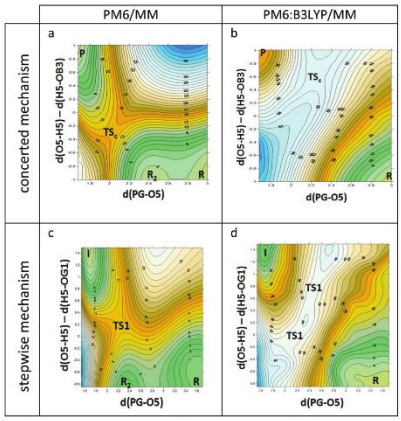


Fig. 3 Free energy surfaces of the concerted and stepwise mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution, obtained as 1D- and 2D-PMFs at PM6/MM level (panels a and c) and with spline corrections at the B3LYP/MM level (panels b and d). Energies are given in kcal mol^{-1} and distances in Å.

According to the coordination of the phosphorous atom at the TSs, these structures cannot be defined as associative TSs since the PG–OB3 distance in both TSs corresponds to a completely broken bond (3.91 and 3.49 Å in AM1d/MM and PM6/MM, respectively). Interestingly, as observed in Fig. 3a, an additional local minimum is located in the reactant valley of the PM6/MM free energy surface that corresponds to reactants with a PG–O5 bond significantly shorter (2.42 Å) than that in reactant structures obtained at AM1d/MM (3.79 Å) and the absolute minimum of reactants obtained with the PM6/MM level (3.11 Å). The corrections of these free energy surfaces at the B3LYP/MM level render two different scenarios. Thus, while the AM1d:B3LYP/MM surface can be considered as quite similar to the original AM1d/MM (the quadratic region of the TS appears at values of the two reaction coordinates very close to those obtained in the original AM1/MM surface), the position of the TS at the PM6:B3LYP/MM surface is significantly shifted with respect to the original position at the PM6/MM one. Thus, according to the corrected surface, the TS would be described by a H5 atom almost completely transferred (O5–H5 and OB3–H5 distance *ca.* to 1.62 and 1.07 Å, respectively), while the forming PG–O5 bond would be in a much earlier stage of the reaction (PG–O5 distance *ca.* 2.3 Å).

From the energetics point of view, the barrier obtained at the AM1d/MM level is much higher than the one deduced from the PM6/MM calculations: 35 and 14 kcal mol^{-1}, respectively. At this point, this huge difference can be related to the

Table 1 QM/MM averaged key distances (in Å) obtained in the states located along the free energy reaction path of the phosphoryl transfer from ATP to Dha in aqueous solution

Distances	PG–O5	PG–OB3	PG–PB	O5–H5	OB3–H5	OG1–H5
AM1d/MM concerted mechanism						
R	3.79 ± 0.03	1.90 ± 0.05	3.37 ± 0.07	0.97 ± 0.03	2.31 ± 0.04	3.43 ± 0.61
TS_C	2.02 ± 0.03	3.91 ± 0.15	4.93 ± 0.18	1.07 ± 0.03	1.57 ± 0.04	3.67 ± 0.10
P	1.78 ± 0.02	4.33 ± 0.33	5.44 ± 0.27	3.19 ± 0.04	0.97 ± 0.03	4.79 ± 0.22
AM1d/MM stepwise mechanism						
R	3.70 ± 0.03	1.71 ± 0.04	3.01 ± 0.07	1.01 ± 0.03	3.74 ± 0.44	1.86 ± 0.04
TS1	2.22 ± 0.03	1.75 ± 0.05	3.03 ± 0.07	2.01 ± 0.04	3.33 ± 0.08	1.01 ± 0.03
I	1.68 ± 0.03	3.82 ± 0.13	4.46 ± 0.24	3.12 ± 0.20	1.88 ± 0.04	1.03 ± 0.03
TS2	1.71 ± 0.04	3.52 ± 0.12	4.43 ± 0.22	3.54 ± 0.16	1.18 ± 0.03	1.34 ± 0.04
P	1.71 ± 0.03	3.54 ± 0.16	4.95 ± 0.18	4.08 ± 0.12	1.04 ± 0.03	1.76 ± 0.04
PM3/MM stepwise mechanism						
R	3.61 ± 0.03	1.77 ± 0.04	3.34 ± 0.05	0.97 ± 0.02	4.12 ± 0.15	1.72 ± 0.03
TS1-1	3.62 ± 0.03	1.74 ± 0.04	3.31 ± 0.06	1.14 ± 0.03	3.53 ± 0.13	1.16 ± 0.03
I_0	3.73 ± 0.03	1.71 ± 0.04	3.27 ± 0.06	1.71 ± 0.03	3.43 ± 0.09	0.97 ± 0.02
TS1-2	2.44 ± 0.03	1.76 ± 0.04	3.33 ± 0.05	1.77 ± 0.03	3.31 ± 0.10	0.96 ± 0.02
I	1.74 ± 0.04	3.86 ± 0.11	5.04 ± 0.16	3.42 ± 0.15	1.77 ± 0.03	0.97 ± 0.02
TS2	1.77 ± 0.04	3.52 ± 0.11	4.82 ± 0.14	3.60 ± 0.16	1.13 ± 0.04	1.16 ± 0.04
P	1.79 ± 0.04	3.75 ± 0.20	5.21 ± 0.15	3.65 ± 0.28	0.97 ± 0.02	1.76 ± 0.03
PM6/MM concerted mechanism						
R	3.11 ± 0.03	1.74 ± 0.05	3.11 ± 0.06	1.02 ± 0.04	2.01 ± 0.05	3.25 ± 0.54
R2	2.42 ± 0.03	1.76 ± 0.05	3.16 ± 0.06	1.01 ± 0.04	2.05 ± 0.05	3.16 ± 0.20
TS_C	2.09 ± 0.03	3.49 ± 0.62	4.69 ± 0.63	1.15 ± 0.03	1.45 ± 0.04	3.62 ± 0.17
P	1.76 ± 0.03	3.72 ± 0.20	4.68 ± 0.24	1.75 ± 0.04	1.05 ± 0.03	4.17 ± 0.15
PM6/MM stepwise mechanism						
R	3.67 ± 0.03	1.78 ± 0.06	3.13 ± 0.07	1.02 ± 0.03	4.14 ± 0.21	1.77 ± 0.04
R2	2.28 ± 0.03	1.93 ± 0.10	3.24 ± 0.09	1.04 ± 0.03	4.17 ± 0.12	1.86 ± 0.04
TS1	2.22 ± 0.03	1.77 ± 0.05	2.84 ± 0.07	1.42 ± 0.04	3.58 ± 0.08	1.24 ± 0.04
I	1.65 ± 0.02	6.66 ± 0.48	6.99 ± 0.85	2.46 ± 0.05	8.21 ± 0.60	1.03 ± 0.04

relative position of the two species, Dha and ATP, in the reactants. As deduced from the interatomic distances reported in Table 1, PM6/MM reactant state structures are described by the two species significantly closer to each other than the reactant state structures located at the AM1d/MM level. This is confirmed by the inspection of PG–O5 and OB3–H5 distances which are smaller in the PM6/MM results (3.11 and 2.01 Å, respectively) than in the AM1d/MM results (3.79 and 2.31 Å, respectively). Then, the TS would be closer to the reactant state in the PM6/MM free energy surface than in the AM1d/MM and, consequently, a smaller barrier is obtained.

When the surfaces are corrected at the B3LYP/MM level, the resulting free energy barriers are 34 and 55 kcal mol^{-1} at AM1d:B3LYP/MM and PM6:B3LYP/MM levels, respectively.

As observed, while the original AM1d/MM surface presents a topology quite close to the corrected AM1d:B3LYP/MM one, which is associated with very close free energy barriers (35 and 34 kcal mol^{-1}, respectively), the fact that corrections of the PM6/MM surface render a dramatically different picture of the topology provokes significantly different barriers (14 and 55 kcal mol^{-1}, respectively).

Stepwise mechanism

As indicated in the Computational methods section, the first step of this mechanism, from reactants to the intermediate, as depicted in Scheme 2, has been explored through the elaboration of a 2D-PMF and the second step, from the intermediate to products, by a 1D-PMF. The resulting free energy surfaces are shown in Fig. 1–3. The reaction pathway of the first step obtained by means of AM1d/MM and PM6/MM levels shows a mechanism where the transfer of the phosphate group and the proton transfer occurs simultaneously leading to the intermediate I. However, despite the advance of the phosphate transfer is equivalent in both cases and located in an early stage of the process (distances PG–O5 equal to 2.22 ± 0.03 Å), the location of the TS on both surfaces is clearly different. Thus, while TS1 in the AM1d/MM surface describes an almost completely transferred proton (distances of O5–H5 and OG1–H5 equal to 2.01 ± 0.04 Å and 1.01 ± 0.03 Å, respectively), the PM6/MM method describes a TS1 with the position of the transferred proton almost in between the donor and acceptor atoms (distances of O5–H5 and OG1–H5 equal to 1.42 ± 0.04 Å and 1.24 ± 0.04 Å, respectively). Moreover, a second minimum corresponding to the reactant state (R2) is obtained in the PM6/MM free energy surfaces (see Fig. 3c), where the two species are in a significant close contact (distance PG–O5 equal to 2.28 ± 0.03 Å). This behavior is similar to the one already detected when exploring the concerted mechanism with this PM6/MM method, where a reactant complex closer to the TS was detected (see Fig. 3a).

The surface obtained at the PM3/MM level is dramatically different from the two surfaces obtained with AM1d/MM and PM6/MM. In this case, an additional minimum corresponding to an intermediate, labeled as I_0 in Fig. 2a, is located. According to the averaged values of the interatomic distances, I_0 corresponds to structures where the proton has been transferred to the OG1 atom (distance OG1–H5 equal to 0.97 ± 0.02 Å) while the phosphate is still anchored to the ADP (distance PG–O5 equal to 3.73 ± 0.03 Å). Then, this first step of the PM3/MM stepwise mechanism is, in fact, a stepwise mechanism controlled by TS1-1 and TS1-2, as labeled in Fig. 2a.

TS1-1 is associated with the proton transfer and TS1-2 with the phosphate transfer. TS1-2 is similar to the TS1 obtained within the AM1d/MM method, as can be confirmed by analysis of key distances listed in Table 1.

When these free energy surfaces are corrected with spline functions at the B3LYP/MM level, interesting observations can be noted. First, the topology of the AM1d:B3LYP/MM surface is almost coincident with the original AM1d/MM, describing a process controlled by TS1 that is basically associated with the phosphate transfer while the proton has been already transferred to the acceptor oxygen atom OG1. On the contrary, the effect of the corrections on the PM3/MM and PM6/MM free energy surfaces is more dramatic. Thus, the minimum corresponding to the I_0 intermediate located on the PM3/MM disappears after HL corrections and, consequently, the step is a concerted process, as obtained with the other two methods. This step is, nevertheless, described by a quite associative TS1, with a very advanced PG–O5 forming bond, and a proton H5 in an early stage of the transfer to the OG1. Finally, the corrections on the PM6/MM surface render a free energy surface where the second reactant minimum (R2 in Fig. 3c) has disappeared and the quadratic region of the TS1 appears as a quite extended region. This region covers structures that range from situations where the proton is in between its donor and acceptor atom to TS structures where the proton has been completely transferred to the OG1 atom (see Fig. 3d).

From the energetics point of view, the free energy barriers of the first step deduced from the AM1d/MM, PM3/MM and PM6/MM methods are 26, 14 and 12 kcal mol^{-1}, respectively. After spline corrections at the B3LYP/MM level, these barriers are transformed into 37, 31 and 45 kcal mol^{-1}, respectively. As observed, there is a clear trend in increasing the values, less dramatic in the case of the AM1d/MM method. From the thermodynamic point of view, PM3/MM and PM6/MM describe a first step of the stepwise mechanism as almost thermoneutral while an endergonic reaction is obtained with the AM1d:B3LYP/MM (8 kcal mol^{-1}).

The free energy surfaces of the second step (from I to P as depicted in Scheme 2) corresponding to the proton transfer from the protonated Dha-P to the ADP to generate the final products, Dha-P and protonated ADP, are shown in Fig. 1 and 2 for the AM1d/MM and PM3/MM results, respectively.

The first observation that must be noticed is that any effort to transfer the proton from OG1 to OB3 did not render any stable conformation when employing the PM6/MM level. When the step was studied with the AM1d/MM and the PM3/MM methodology, barriers of 7 kcal mol^{-1} and 18 kcal mol^{-1}, respectively, were obtained.

Analysis of the key interatomic distances reported in Table 1 reveals an AM1d/MM TS2 with a more advanced proton transfer (H5–OG1 and H5–OB3 distances of TS2 at the

AM1d/MM level are 1.34 ± 0.04 Å and 1.18 ± 0.03 Å, respectively) than in the PM3/MM TS2 (H5–OG1 and H5–OB3 distances of TS2 at the PM3/MM level are 1.16 ± 0.04 Å and 1.13 ± 0.04 Å, respectively).

Moreover, by comparing I with P, more stabilized products are achieved from PM3/MM calculations than from the AM1d/MM calculations (2 kcal mol^{-1} and 6 kcal mol^{-1}, respectively). The corrected free energy profiles show a slight increase of the barrier in the case of AM1d:B3LYP/MM (*ca.* 5 kcal mol^{-1}) and, in contrast, a slight decrease when using PM3:B3LYP/MM (*ca.* 3 kcal mol^{-1}).

The overall free energy profiles for concerted and stepwise mechanisms obtained at the different levels of calculations are compared in Fig. 4 while representative snapshots of struc-

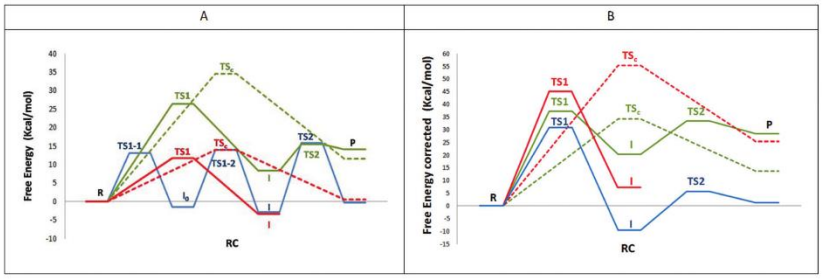


Fig. 4 QM/MM free energy profiles of the two possible reaction mechanisms corresponding to the phosphoryl transfer from ATP to Dha: the stepwise mechanism (solid lines) and the concerted mechanism (dashed lines). A panel displays the results from QM/MM calculations with the QM region treated by AM1d (green line), PM3 (blue line) and PM6 (red line). B panel displays the results from AM1d:B3LYP/MM (green line), PM3:B3LYP/MM (blue line) and PM6:B3LYP/MM (red line) methods.

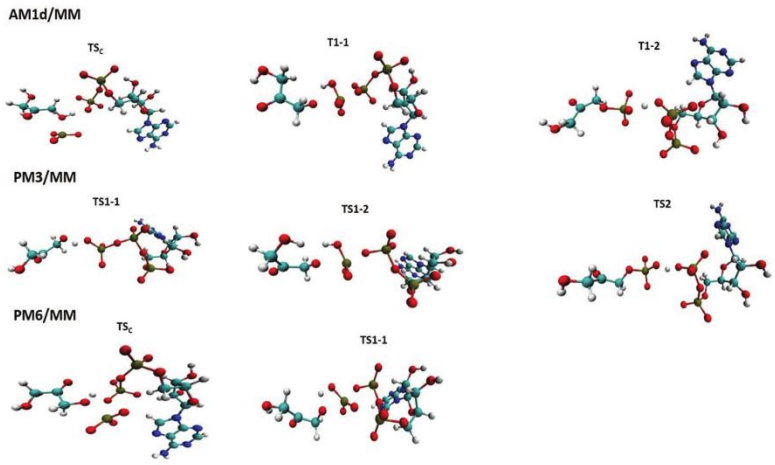


Fig. 5 Representative snapshots of structures of the TSs obtained with the different methods. Water molecules are not displayed for clarity purposes. TS_C corresponds to the TS of the concerted mechanism while TS1-1, TS1-2, and TS2 corresponds to the different TSs along the stepwise mechanism.

tures of the located TSs are depicted in Fig. 5. The corresponding relative energies to reactants are listed in Table S1 of the ESI.† As observed in Fig. 4, in the AM1d/MM method, the stepwise mechanism is favored over the concerted one, but the two mechanisms are very endergonic. The free energy profiles obtained with PM6/MM present lower barriers but with this methodology, it is impossible to localize the final products through the stepwise mechanism. PM3/MM also shows very low free energy barriers, compared with AM1d/MM, to obtain the intermediates and the final product, achieving a high stabilization of both. Keeping in mind the low barriers obtained at the PM3/MM level, an alternative possible solvent-assisted mechanism was explored for the rate limiting step of this profile. Nevertheless, the energy barrier was 29 kcal mol^{-1}, once again much higher than the corresponding step without the participation of a quantum water molecule (see PES in Fig. S4 in the ESI†).

Free energy profiles obtained after B3LYP/MM corrections (see panel B of Fig. 4) show how the highest barriers are obtained with the PM6:B3LYP/MM method. Free energy barriers of the stepwise mechanism obtained with AM1d:B3LYP/MM and PM3:B3LYP/MM are quite similar for the first step, but much smaller with the latter method.

Conclusions

In this paper, we have studied the phosphoryl transfer reaction between adenosine triphosphate (ATP) and Dha in aqueous solution within different hybrid QM/MM potentials. In particular, different semiempirical Hamiltonians have been employed to describe the sub-set of atoms treated quantum mechanically, that basically contains the two involved species; ATP and Dha. A deep understanding of the reaction, which is considered as ubiquitous in biology due to the relevance of the formed products, as well as the comparative analysis between different methods, can have implications for further studies of the catalyzed reaction by the action of dihydroxyacetone kinase.

The first conclusion that can be derived from the calculations is that the possible solvent-assisted mechanisms render significant higher potential energy barriers than the corresponding steps without explicit participation of a water molecule. This possibility was then discarded and attention was focused in exploring the free energy surfaces of the rest of the proposed reaction paths. After exploring the free energy surfaces obtained in terms of 1D and 2D-PMFs for the concerted and stepwise mechanisms at AM1d/MM, PM3/MM and PM6/MM levels, interesting conclusions can be derived. First of all, we have demonstrated the high dependency of the results on the selected Hamiltonian. Thus, while AM1d/MM and PM6/MM methods can describe a concerted mechanism, any attempt to obtain this mechanism with PM3/MM was unsuccessful. Regarding the stepwise mechanism, AM1d/MM and PM3/MM were able to describe both steps, phosphate transfer and hydride transfer but, in the case of the PM6/MM method, the second step was not located, the intermediate (protonated Dha-P) being much more stable than the products (protonated ADP). A comparison of the topology of the free energy surfaces of the concerted mechanism obtained with AM1d/MM and PM6/MM methods shows a similar location of the quadratic region of the TS, defined by a phosphate transfer in an advanced stage of the process but a proton transfer still bound to the Dha moiety. On the contrary, regarding the stepwise mechanism, the location of the TS of the first step obtained with AM1d/MM is quite different from the one obtained with PM6/MM (the phosphate group is transferred after protonation from Dha according to the AM1d/MM method while the opposite timing is obtained with the PM6/MM). PM3/MM renders a surface that describes this step in two-steps; first the transfer of the proton from Dha followed by the protonated phosphate group transfer. This second step is, in fact, similar to the TS1 located with AM1d/MM. The last step of the stepwise mechanism obtained with AM1d/MM and PM3/MM is quite similar regarding the TS, although differences are obtained regarding the thermodynamics of the reaction.

When combining all energetic data of every single step, the full free energy profiles (depicted in Fig. 4) show that the stepwise mechanism appears as more favorable than the concerted one when atoms included in the QM region of our QM/MM scheme are described with semiempirical Hamiltonians, being PM3/MM the method rendering the lowest free energy barriers for this stepwise mechanism. This result is reasonable, keeping in mind that the TS of the concerted mechanism requires a non-inline attack (see Fig. 5), associated with a higher energy transient conformation. Moreover, the reaction described by the PM3/MM method is almost thermoneutral, while AM1d/MM and PM6/MM methods show endergonic processes. At this point, we must keep in mind that the different products, obtained with the different methods and by exploring the different mechanisms, correspond to different product complexes that can differ in the particular conformations and stabilizing inter-molecular interactions. If the reaction was explored up to the completely solvent separated species, the reaction energy should not depend on the mechanism.

Keeping in mind the limitations of the semiempirical methods to describe this kind of reaction that involves phosphor atoms, spline corrections at the B3LYP/6-31G(d,p)/MM level have been applied to the previous free energy surfaces. The topology of the resulting high level AM1d:B3LYP/MM free energy surfaces is similar to the previously obtained AM1d/MM ones. The free energy barriers are, nevertheless, slightly raised. On the contrary, high level corrections have significant effects on the PM3/MM surfaces, where the two steps located on the first step of the stepwise mechanism are transformed into a concerted one, and the position of the quadratic region of the TS1 is shifted, becoming closer to the position of the AM1d/MM TS1. The B3LYP/MM correction slightly decreases the free energy barrier. Finally, the effects of the corrections on the PM6/MM surfaces of the concerted mechanism and the first step of the stepwise mechanism are also noticeable. The TS of the concerted mechanism describes a reaction where the

proton transfer takes place prior to the phosphate transfer, while the original TS1 is transformed into a flat and wide quadratic region in the corrected surface.

After this analysis, it appears quite obvious the dependency of the resulting mechanism on the employed method. Moreover, it seems that corrections at the B3LYP/MM level render different results depending on the original structures generated during the exploration of the 1D and 2D-PMFs. Thus, in order to look for a convincing proof of the most trustworthy reaction path, PESs have been generated at the B3LYP/MM level of theory. We must keep in mind that B3LYP/MM MD calculations are prohibitive from the computational point of view. The resulting surfaces are reported in Fig. 6, while structures of the located TSs are displayed in Fig. 7.

As observed, and by comparison with the free energy surfaces displayed in Fig. 1–3, and the corresponding PESs reported in the ESI,† the results obtained at AM1d/MM are the ones that resemble the most to the B3LYP/MM PESs. This explains the small effect on the topology of the free energy surfaces after B3LYP/MM corrections. Thus, we can confirm the re-parametrized AM1d Hamiltonian as the most convenient semiempirical method to be used in QM/MM MD simulations to explore the reactivity of phosphoryl transfer reactions in terms of free energy surfaces.

The fact that products appear as a shallow minimum in the B3LYP/MM PES of the concerted mechanism (see Fig. 6a) could be considered in agreement with the fact that no concerted mechanism was obtained at the PM3/MM level. Moreover, as observed in Fig. 6c, the product complex is less stable than the intermediate in the stepwise mechanism and shows an almost negligible backward barrier to transform into the intermediate. This could be also considered in agreement with the fact that this second step was not located at the PM6/MM level.

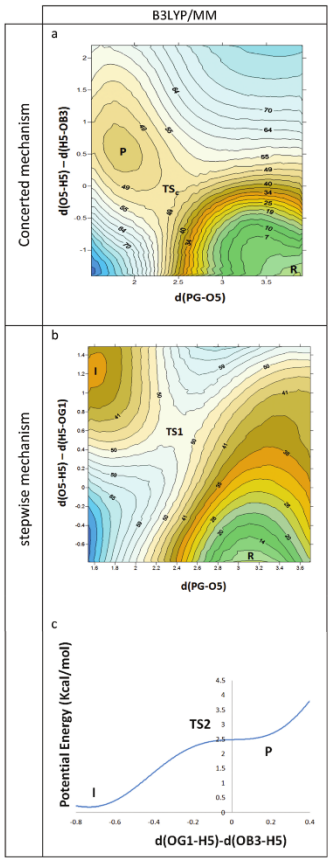


Fig. 6 Potential energy surfaces of the concerted (a) and stepwise mechanisms (b and c) of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at the B3LYP/MM level. Energies are reported in kcal mol^{-1} and distances in Å.

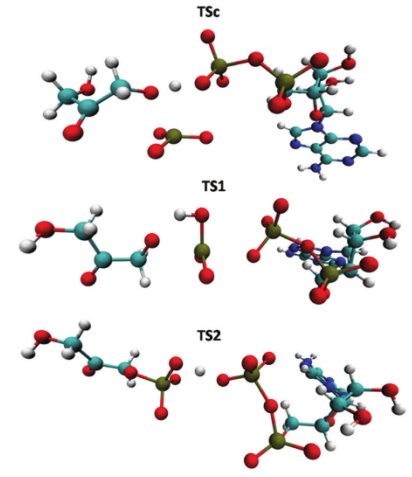


Fig. 7 Representative snapshots of structures of the TSs obtained with the B3LYP/MM method. Water molecules are not displayed for clarity purposes.

As mentioned, the result of the comparative analysis between different hybrid QM/MM methods applied to the phosphoryl transfer reaction between ATP and Dha in aqueous solution can have implications for further studies of the catalyzed reaction by the action of dihydroxyacetone kinase. Nevertheless, keeping in mind the features of the reaction, basically, the highly charged involved species, dramatic differences can be expected when including the effect of the protein environment. In particular, positively charged residues and Mg^{2+} ions, that are present in the active site of the kinases, can have an important role in stabilization of transient species. From the technical point of view, these residues should most probably be included in the QM region during the QM/MM because of the presumably large charge reorganization. The analysis of the timescale of protein motions involving these residues will allow having a dynamic description of the process. This prediction is based on the previous comparative analysis of chemical reactivity in protein and aqueous environments. Finally, it is important to point out that this study can shed some light in the debate about the kinds of TSs governing similar reactions involving phosphate groups.

Acknowledgements

This work was supported by the Spanish Ministerio de Economia y Competitividad (project CTQ2012-36253-C03-01), Generalitat Valenciana (project Prometeo II/2014/022) and Universitat Jaume I (project P1 1B2013-58). The authors acknowledge computational facilities of the Servei d'Informàtica of Universitat Jaume I.

Notes and references

1 *Phosphoryl Transfer Reactions*, ed. A. C. Hengge, John Wiley & Sons, 2001.
2 S. C. L. Kamerlin, N. H. Williams and A. Warshel, *J. Org. Chem.*, 2008, **73**, 6960.
3 S. C. L. Kamerlin, *J. Org. Chem.*, 2011, **76**, 9228.
4 C. Bachler, K. Flukiger-Bruhwiler, P. Schneider, P. Bahler and B. Erni, *J. Biol. Chem.*, 2005, **280**, 18321.
5 S. Iwamoto, K. Motomura, Y. Shinoda, M. Urata, J. Kato, N. Takiguchi, H. Ohtake, R. Hirota and A. Kuroda, *Appl. Environ. Microbiol.*, 2007, **73**, 5676.
6 B. Erni, C. Siebold, S. Christen, A. Srinivas, A. Oberholzer and U. Baumann, *Cell. Mol. Life Sci.*, 2006, **63**, 890.
7 R. Shi, L. McDonald, Q. Cui, A. Matte, M. Cygler and I. Ekiel, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 1302.
8 D. Enders, M. Voith and A. Lenzen, *Angew. Chem., Int. Ed.*, 2005, **44**, 1304.
9 J. Colinge, A. Cesar-Razquin, K. Huber, F. P. Breitwieser, P. Majek and G. Superti-Furga, *J. Proteomics*, 2014, **107**, 113.
10 H. Weinmann and R. Metternich, *ChemBioChem*, 2005, **6**, 455.
11 J. K. Lassila, J. G. Zalatan and D. Herschlag, in *Annual Review of Biochemistry*, ed. R. D. Kornberg, C. R. H. Raetz, J. E. Rothman and J. W. Thorner, 2011, vol. 80, p. 669.
12 R. P. Bora, M. J. L. Mills, M. P. Frushicheva and A. Warshel, *J. Phys. Chem. B*, 2015, **119**, 3434.
13 B. L. Grigorenko, A. V. Rogov and A. V. Nemukhin, *J. Phys. Chem. B*, 2006, **110**, 4407.
14 C. B. Harrison and K. Schulten, *J. Chem. Theory Comput.*, 2012, **8**, 2328.
15 F. Xia, K. Tian and H. Zhu, *J. Chem. Theory Comput.*, 2013, **1017**, 60.
16 N. V. Plotnikov, B. R. Prasad, S. Chakrabarty, Z. T. Chu and A. Warshel, *J. Phys. Chem. B*, 2013, **117**, 12807.
17 J. Florian and A. Warshel, *J. Phys. Chem. B*, 1998, **102**, 719.
18 S. C. Kamerlin, P. K. Sharma, R. B. Prasad and A. Warshel, *Q. Rev. Biophys.*, 2013, **46**, 1.
19 C. H. Hu and T. Brinck, *J. Phys. Chem. A*, 1999, **103**, 5379.
20 F. Duarte, J. Aqvist, N. H. Williams and S. C. L. Kamerlin, *J. Am. Chem. Soc.*, 2015, **137**, 1081–1093.
21 Y. Liu, X. Lopez and D. M. York, *Chem. Commun.*, 2005, 3909.
22 X. Lopez, A. Dejaegere, F. Leclerc, D. M. York and M. Karplus, *J. Phys. Chem. B*, 2006, **110**, 11525.
23 D G. Xu, H. Guo, Y. Liu and D. M. York, *J. Phys. Chem. B*, 2005, **109**, 13827.
24 K. Nam, Q. Cui, J. Gao and D. M. York, *J. Chem. Theory Comput.*, 2007, **3**, 486.
25 V. Lopez-Canut, S. Marti, J. Bertran, V. Moliner and I. Tuñón, *J. Phys. Chem. B*, 2009, **113**, 7816.
26 V. Lopez-Canut, J. Javier Ruiz-Pernia, R. Castillo, V. Moliner and I. Tuñón, *Chem. – Eur. J.*, 2012, **18**, 9612.
27 V. Lopez-Canut, M. Roca, J. Bertran, V. Moliner and I. Tuñón, *J. Am. Chem. Soc.*, 2011, **133**, 12050.
28 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209.
29 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215.
30 J. J. P. Stewart, *Journal of Molecular Modeling*, 2007, **13**, 1173.
31 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700.
32 W. L. Jorgensen and J. Tiradorives, *J. Am. Chem. Soc.*, 1988, **110**, 1657.
33 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
34 M. J. Field, M. Albe, C. Bret, F. Proust-De Martin and A. Thomas, *J. Comput. Chem.*, 2000, **21**, 1088.
35 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, 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, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A. 2*, Gaussian; Inc, Wallingford, CT, 2009.
36 B. Roux, *Comput. Phys. Commun.*, 1995, **91**, 275.
37 G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187.
38 S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman and J. M. Rosenberg, *J. Comput. Chem.*, 1992, **13**, 1011.
39 J. C. Corchado, E. L. Coitino, Y. Y. Chuang, P. L. Fast and D. G. Truhlar, *J. Phys. Chem. A*, 1998, **102**, 2424.
40 K. A. Nguyen, I. Rossi and D. G. Truhlar, *J. Chem. Phys.*, 1995, **103**, 5522.
41 Y. Y. Chuang, J. C. Corchado and D. G. Truhlar, *J. Phys. Chem. A*, 1999, **103**, 1140.
42 R. J. Renka, *SIAM J. Sci. Stat. Comput.*, 1987, **8**, 393.
43 R. J. Renka, *ACM Trans. Math. Software*, 1993, **19**, 81.
44 J. J. Ruiz-Pernia, E. Silla, I. Tuñon and S. Marti, *J. Phys. Chem. B*, 2006, **110**, 17663.
45 S. Ferrer, S. Marti, V. Moliner, I. Tuñón and J. Bertran, *Phys. Chem. Chem. Phys.*, 2012, **14**, 3482.

Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supporting Information

A Computational Study of the Phosphoryl Transfer Reaction between ATP and Dha in Aqueous Solution

I. Bordes, J. J. Ruiz-Pernía, R. Castillo* and V. Moliner*

Departament de Química Física i Analítica, Universitat Jaume I,
12071Castellón (Spain)

* to whom correspondence should be addressed:
R. Castillo: rcastill@uji.es
V. Moliner: molinerr@uji.es

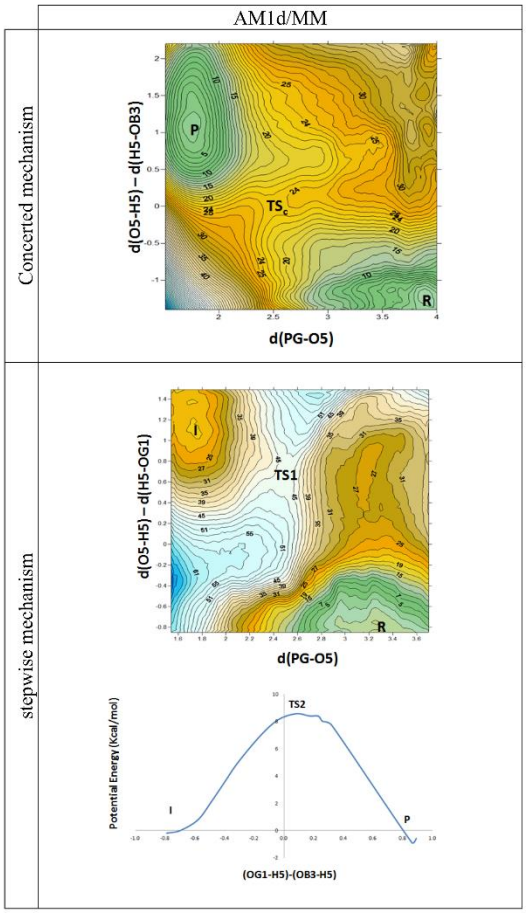


Figure S1. PES of the concerted and stepwise mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at AM1d/MM level. Energies are given in kcal·mol^{-1} and distances in Å.

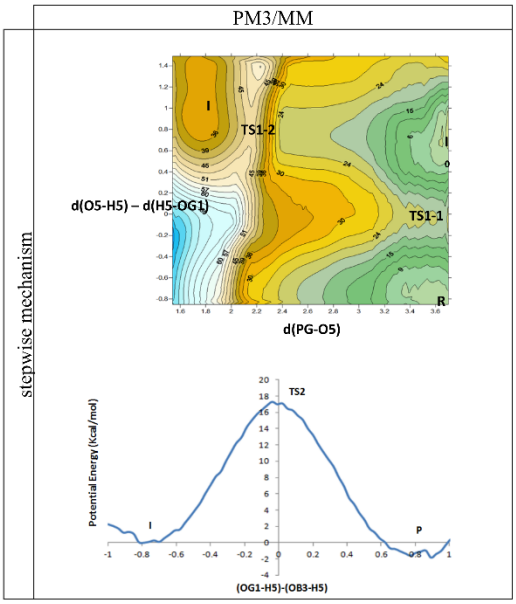


Figure S2. PES and potential energy profile of the stepwise mechanism of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at PM3/MM level. Energies are given in kcal·mol^{-1} and distances in Å.

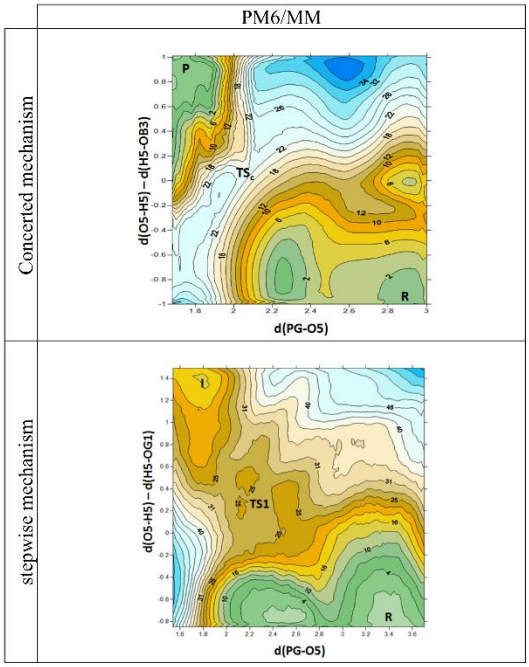


Figure S3. PES of the concerted and stepwise mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at PM6/MM level. Energies are given in kcal·mol^{-1} and distances in Å.

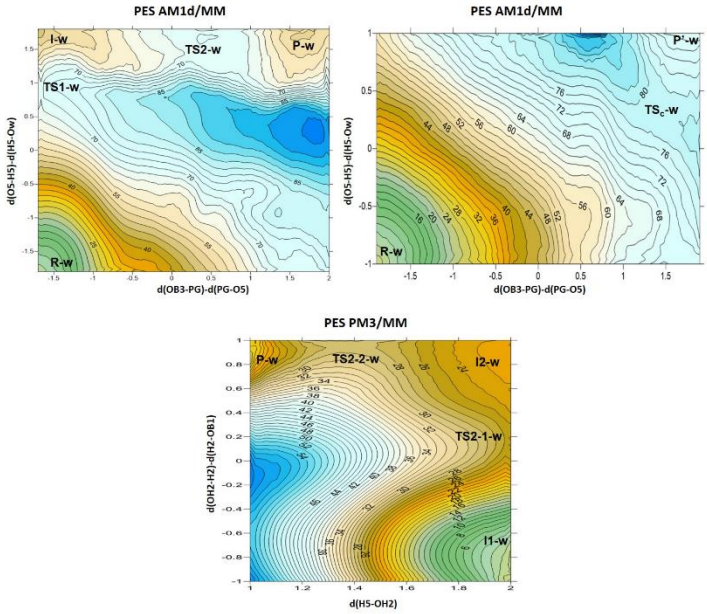


Figure S4. PES of the stepwise solvent-assisted mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at AM1d/MM level for the concerted and stepwise mechanisms (on the top) and PM3/MM level (down part). Energies are given in kcal·mol^{-1} and distances in Å.

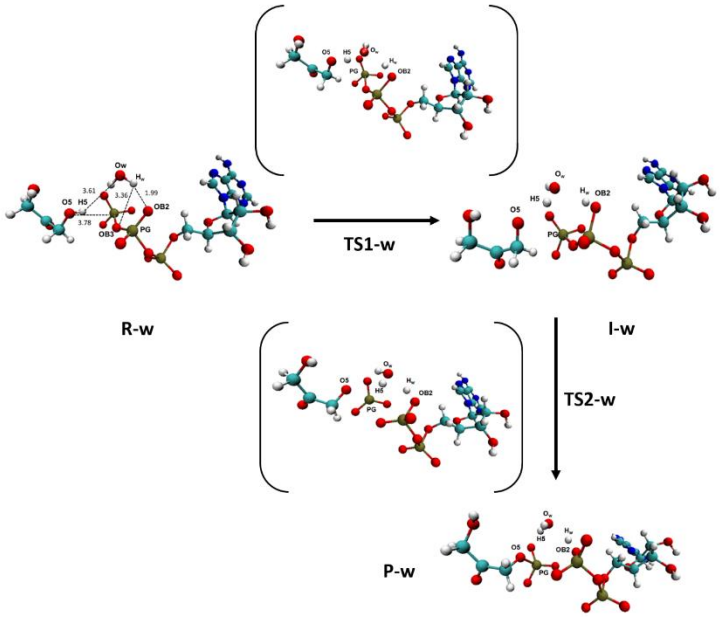


Figure S5. Geometries of the stationary points of the stepwise solvent-assisted mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at AM1d/MM level. Distances in Å.

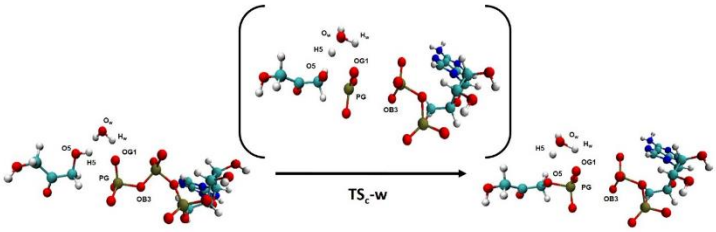


Figure S6. Geometries of the stationary points of the concerted solvent-assisted mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at AM1d/MM level. It must be pointed out that **P'-w** correspond to a stationary state previous to the final **P-w**.

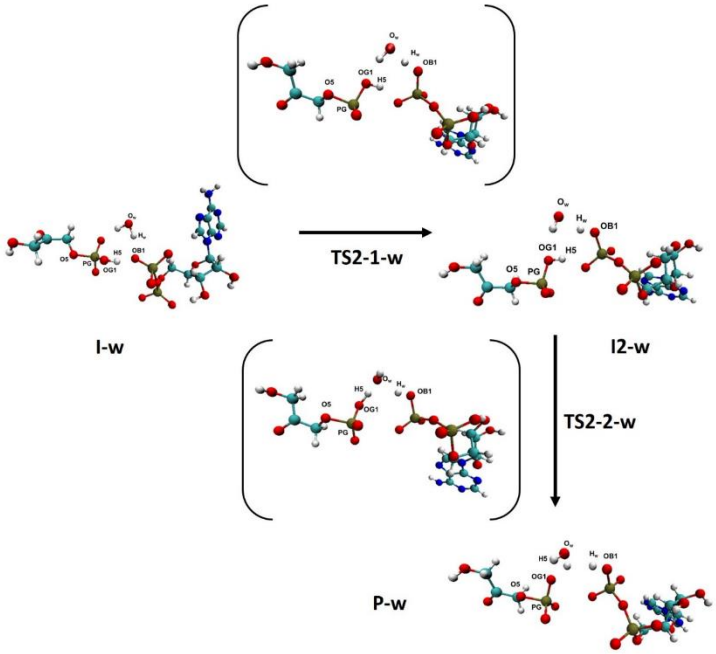


Figure S7. Geometries of the stationary points of the stepwise solvent-assisted mechanisms of the phosphoryl transfer reaction from ATP to Dha in aqueous solution obtained at PM3/MM level.

Table S1. QM/MM free energies relative to reactants of the two possible reaction mechanisms corresponding to the phosphoryl transfer from ATP to Dha. Table shows the results from QM/MM calculations with the QM region treated by AM1d/MM, PM3/MM and PM6/MM. Energies are given in kcal·mol^{-1}.

AM1d/MM		**PM3/MM**		**PM6/MM**	
Concerted mechanism		Stepwise mechanism		Concerted mechanism	
R	0.00	**R**	0.00	**R**	0.00
TS$_c$	34.59	**TS1-1**	13.24	**TS$_c$**	14.0
P	11.63	**I$_0$**	-1.45	**P**	0.51
Stepwise mechanism		**TS1-2**	14.09	Stepwise mechanism	
R	0.00	**I (TS1-2)**	-2.82	**R**	0.00
TS1	26.40	**I (TS2)**	-2.82	**TS1**	11.69
I (TS1)	8.29	**TS2**	15.91	**I**	-3.39
I (TS2)	8.29	**P**	-0.26		
TS2	15.49				
P	14.13				

Table S2. QM/MM free energies relative to reactants of the two possible reaction mechanisms corresponding to the phosphoryl transfer from ATP to Dha. Table shows the results from QM/MM calculations with the QM region treated by AM1d:B3LYP/MM, PM3:B3LYP/MM and PM6:B3LYP/MM. Energies are given in kcal·mol^{-1}.

AM1d:B3LYP/MM		**PM3:B3LYP/MM**		**PM6:B3LYP/MM**	
Concerted mechanism		Stepwise mechanism		Concerted mechanism	
R	0.00	**R**	0.00	**R**	0.00
TS$_c$	34.22	**TS1**	30.92	**TS$_c$**	55.32
P	13.62	**I (TS1)**	-9.70	**P**	25.42
Stepwise mechanism		**I (TS2)**	-9.70	Stepwise mechanism	
R	0.00	**TS2**	5.75	**R**	0.00
TS1	37.28	**P**	1.25	**TS1**	45.01
I	20.36			**I**	7.22
Iprofile	20.36				
TS2	33.44				
P	28.51				

Tables S4: Cartesian coordinates of the QM atoms of the localized stationary points.

AM1d/MM Concerted mechanism	R			TS_c			P		
C1 DHA	-7.599	0.988	-0.198	-7.248	0.940	-0.229	-7.174	0.875	-0.172
C2 DHA	-6.364	1.796	-0.459	-6.100	1.865	-0.590	-5.916	1.683	-0.475
H21 DHA	-6.650	2.843	-0.777	-6.360	2.346	-1.578	-6.107	2.268	-1.420
H22 DHA	-5.689	1.834	0.459	-6.073	2.658	0.212	-5.767	2.399	0.383
C3 DHA	-8.423	0.359	-1.289	-8.170	0.423	-1.319	-8.151	0.510	-1.277
H31 DHA	-9.117	-0.428	-0.874	-8.933	-0.274	-0.871	-9.010	-0.072	-0.835
H32 DHA	-9.014	1.164	-1.816	-8.701	1.306	-1.770	-8.542	1.465	-1.723
O4 DHA	-8.032	0.879	1.004	-7.583	0.738	0.948	-7.523	0.634	0.994
O5 DHA	-5.643	1.243	-1.572	-4.849	1.267	-0.722	-4.764	0.924	-0.705
H5 DHA	-4.871	0.692	-1.227	-3.995	2.254	-0.670	-2.630	2.575	0.540
O6 DHA	-7.605	-0.338	-2.240	-7.490	-0.326	-2.296	-7.581	-0.328	-2.256
H6 DHA	-7.295	0.307	-2.959	-7.290	0.270	-3.049	-7.395	0.224	-3.043
C1' ATP	5.252	1.568	0.113	5.209	1.531	0.075	5.253	1.546	0.124
H1' ATP	5.987	2.255	0.635	5.984	2.194	0.559	6.041	2.199	0.605
C2' ATP	5.167	1.934	-1.381	5.152	1.903	-1.418	5.230	1.865	-1.383
H2' ATP	4.803	1.100	-2.048	4.750	1.074	-2.061	4.868	1.001	-2.003
O2' ATP	6.470	2.229	-1.894	6.434	2.172	-1.918	6.521	2.143	-1.858
HT2 ATP	6.865	3.035	-1.443	6.807	2.937	-1.440	6.858	2.940	-1.408
C3' ATP	4.186	3.127	-1.401	4.182	3.104	-1.426	4.236	3.042	-1.480
H3' ATP	3.611	3.212	-2.366	3.548	3.126	-2.349	3.588	2.962	-2.390
O3' ATP	4.968	4.330	-1.232	4.957	4.282	-1.313	4.992	4.236	-1.518
HT3 ATP	4.390	5.116	-1.430	4.401	5.019	-1.610	4.368	4.957	-1.695
C4' ATP	3.296	2.927	-0.155	3.342	2.916	-0.145	3.408	2.948	-0.182
O4' ATP	3.942	1.888	0.669	3.944	1.867	0.647	3.994	1.940	0.671
H4' ATP	3.273	3.841	0.509	3.370	3.849	0.481	3.455	3.910	0.391
C5' ATP	1.884	2.469	-0.508	1.902	2.504	-0.431	1.965	2.527	-0.439
H51 ATP	1.790	1.358	-0.664	1.839	1.401	-0.643	1.901	1.418	-0.610
H52 ATP	1.493	3.012	-1.419	1.513	3.061	-1.327	1.570	3.052	-1.352
O5' ATP	1.043	2.740	0.651	1.136	2.748	0.730	1.201	2.805	0.718
PA ATP	0.033	4.078	0.663	0.045	4.218	0.719	0.153	4.299	0.689
OA1 ATP	0.897	5.286	0.349	0.997	5.437	0.455	1.101	5.529	0.492
OA2 ATP	-0.753	3.976	1.945	-0.875	4.032	1.964	-0.822	4.064	1.893
OA3 ATP	-0.896	3.860	-0.639	-0.757	4.073	-0.812	-0.588	4.171	-0.894
PB ATP	-1.777	2.572	-1.228	-1.826	2.759	-1.482	-1.671	2.864	-1.535
OB1 ATP	-2.466	3.321	-2.362	-2.369	3.580	-2.717	-2.330	3.499	-2.804
OB2 ATP	-0.826	1.468	-1.651	-0.923	1.498	-1.748	-0.847	1.537	-1.647
OB3 ATP	-2.807	2.187	-0.069	-3.023	2.724	-0.306	-2.844	3.086	-0.259
PG ATP	-3.120	0.672	0.683	-3.627	0.201	0.838	-3.825	0.163	0.744
OG1 ATP	-3.545	-0.286	-0.458	-3.861	-1.022	-0.133	-4.010	-1.301	0.127
OG2 ATP	-1.813	0.280	1.371	-2.096	0.417	1.120	-2.279	0.573	0.826
OG3 ATP	-4.280	0.978	1.621	-4.670	0.704	1.870	-4.660	0.650	1.991

AM1d/MM Stepwise mechanism	TS1	I	TS2
C1 DHA	-7.263 -0.755 0.837	-7.387 -0.818 0.841	-7.402 -0.852 0.840
C2 DHA	-5.771 -0.546 0.799	-5.881 -0.724 0.855	-5.905 -0.813 0.858
H21 DHA	-5.401 -0.439 1.861	-5.530 -0.805 1.924	-5.536 -0.731 1.926
H22 DHA	-5.332 -1.472 0.319	-5.487 -1.615 0.282	-5.505 -1.753 0.363
C3 DHA	-8.143 -0.109 -0.182	-8.174 -0.042 -0.175	-8.199 -0.106 -0.174
H31 DHA	-8.092 1.013 -0.091	-8.030 1.057 0.007	-8.088 1.008 -0.031
H32 DHA	-7.835 -0.436 -1.217	-7.808 -0.303 -1.206	-7.879 -0.402 -1.218
O4 DHA	-7.790 -1.535 1.723	-7.961 -1.577 1.646	-7.989 -1.603 1.707
O5 DHA	-5.392 0.579 0.056	-5.459 0.467 0.265	-5.433 0.315 0.077
H5 DHA	-4.593 1.701 -1.204	-4.740 1.898 -1.255	-2.264 2.091 -0.193
O6 DHA	-9.512 -0.554 -0.077	-9.538 -0.406 -0.156	-9.587 -0.486 -0.121
H6 DHA	-9.937 -0.153 0.722	-9.943 -0.006 0.626	-10.039 -0.054 0.648
C1' ATP	6.187 2.151 -0.333	6.211 2.141 -0.276	6.204 2.150 -0.330
H1' ATP	7.235 2.583 -0.309	7.273 2.528 -0.279	7.258 2.570 -0.321
C2' ATP	5.325 2.907 -1.365	5.390 2.954 -1.297	5.338 2.926 -1.344
H2' ATP	4.996 2.282 -2.244	4.966 2.322 -2.121	4.977 2.309 -2.216
O2' ATP	6.123 3.921 -1.996	6.213 3.885 -1.963	6.148 3.923 -1.988
HT2 ATP	6.134 4.751 -1.450	6.435 4.598 -1.346	6.177 4.755 -1.446
C3' ATP	4.128 3.486 -0.576	4.272 3.620 -0.463	4.171 3.530 -0.531
H3' ATP	3.131 3.132 -0.981	3.247 3.356 -0.845	3.156 3.208 -0.926
O3' ATP	4.183 4.916 -0.746	4.487 5.013 -0.590	4.263 4.961 -0.689
HT3 ATP	3.391 5.317 -0.265	3.874 5.459 0.012	3.525 5.387 -0.146
C4' ATP	4.350 3.114 0.906	4.470 3.129 0.988	4.408 3.139 0.944
O4' ATP	5.650 2.445 0.989	5.708 2.400 1.037	5.695 2.444 1.002
H4' ATP	4.467 4.025 1.568	4.594 3.998 1.692	4.550 4.041 1.614
C5' ATP	3.313 2.146 1.494	3.376 2.190 1.527	3.365 2.184 1.543
H51 ATP	3.778 1.541 2.324	3.854 1.467 2.247	3.825 1.582 2.377
H52 ATP	2.830 1.485 0.721	2.900 1.616 0.686	2.874 1.521 0.776
O5' ATP	2.266 2.883 2.178	2.410 2.877 2.284	2.323 2.935 2.223
PA ATP	1.209 3.853 1.307	1.243 3.960 1.376	1.340 3.976 1.354
OA1 ATP	1.980 5.072 0.818	2.096 5.203 0.889	2.130 5.227 0.990
OA2 ATP	0.058 4.133 2.239	0.055 4.267 2.360	0.120 4.200 2.213
OA3 ATP	0.895 2.794 0.135	0.937 2.857 0.118	1.191 3.067 0.028
PB ATP	-0.471 2.304 -0.721	-0.578 2.252 -0.802	-0.152 2.472 -0.772
OB1 ATP	-0.921 3.548 -1.470	-1.033 3.557 -1.589	-0.701 3.647 -1.572
OB2 ATP	0.135 1.147 -1.514	0.135 1.083 -1.617	0.401 1.254 -1.507
OB3 ATP	-1.487 1.807 0.374	-1.582 1.707 0.346	-1.169 2.089 0.356
PG ATP	-3.209 1.165 0.396	-3.657 0.949 0.512	-3.865 0.808 0.451
OG1 ATP	-3.583 1.714 -1.083	-3.834 1.579 -1.152	-3.381 1.637 -0.791
OG2 ATP	-2.755 -0.273 0.445	-2.902 -0.434 0.586	-3.014 -0.462 0.580
OG3 ATP	-3.621 1.980 1.588	-3.826 2.005 1.669	-3.957 1.703 1.687

PM3/MM Stepwise mechanism	R			TS1-1			I_0			TS1-2		
C1 DHA	-7.895	-0.842	0.948	-7.854	-0.838	0.958	-7.916	-0.874	0.950	-7.562	-0.871	0.973
C2 DHA	-6.437	-0.424	0.907	-6.389	-0.425	1.017	-6.471	-0.366	0.985	-6.038	-0.772	1.010
H21 DHA	-5.918	-0.677	1.853	-5.969	-0.662	2.015	-6.003	-0.666	1.949	-5.670	-1.010	2.028
H22 DHA	-5.934	-0.995	0.099	-5.858	-1.068	0.284	-5.952	-0.950	0.191	-5.662	-1.564	0.330
C3 DHA	-8.839	-0.197	-0.044	-8.710	-0.234	-0.133	-8.798	-0.323	-0.147	-8.322	-0.100	-0.087
H31 DHA	-8.856	0.900	0.124	-8.618	0.871	-0.111	-8.736	0.786	-0.138	-8.137	0.984	0.048
H32 DHA	-8.472	-0.379	-1.071	-8.340	-0.582	-1.114	-8.416	-0.670	-1.125	-7.955	-0.384	-1.090
O4 DHA	-8.277	-1.695	1.748	-8.308	-1.670	1.748	-8.336	-1.707	1.759	-8.148	-1.619	1.759
O5 DHA	-6.356	0.958	0.695	-6.217	0.915	0.752	-6.393	0.958	0.789	-5.566	0.460	0.618
H5 DHA	-5.517	1.118	0.201	-5.255	1.071	0.156	-4.943	1.236	-0.054	-5.172	1.344	-0.935
O6 DHA	-10.133	-0.739	-0.029	-10.055	-0.639	-0.098	-10.133	-0.765	-0.113	-9.696	-0.391	-0.131
H6 DHA	-10.562	-0.413	0.753	-10.467	-0.163	0.613	-10.531	-0.374	0.656	-10.102	0.059	0.601
C1' ATP	6.279	2.243	-0.250	6.275	2.240	-0.246	6.287	2.231	-0.221	6.244	2.225	-0.271
H1' ATP	7.323	2.606	-0.133	7.318	2.604	-0.121	7.334	2.580	-0.085	7.283	2.601	-0.150
C2' ATP	5.531	3.110	-1.287	5.530	3.106	-1.285	5.567	3.106	-1.270	5.485	3.103	-1.291
H2' ATP	5.245	2.530	-2.192	5.249	2.524	-2.190	5.300	2.529	-2.183	5.158	2.524	-2.182
O2' ATP	6.335	4.111	-1.858	6.334	4.108	-1.853	6.383	4.106	-1.822	6.298	4.079	-1.894
HT2 ATP	6.589	4.706	-1.163	6.591	4.698	-1.154	6.650	4.683	-1.116	6.573	4.682	-1.213
C3' ATP	4.277	3.618	-0.523	4.271	3.610	-0.527	4.299	3.610	-0.531	4.265	3.648	-0.501
H3' ATP	3.366	3.124	-0.939	3.365	3.104	-0.940	3.398	3.114	-0.967	3.325	3.214	-0.920
O3' ATP	4.131	5.009	-0.670	4.116	4.997	-0.688	4.158	5.000	-0.678	4.186	5.048	-0.601
HT3 ATP	3.355	5.224	-0.128	3.337	5.214	-0.153	3.367	5.216	-0.164	3.409	5.276	-0.064
C4' ATP	4.496	3.207	0.964	4.491	3.212	0.964	4.487	3.193	0.959	4.477	3.187	0.971
O4' ATP	5.706	2.452	1.042	5.697	2.449	1.043	5.696	2.443	1.062	5.671	2.404	1.026
H4' ATP	4.678	4.091	1.617	4.676	4.102	1.607	4.653	4.072	1.623	4.675	4.045	1.654
C5' ATP	3.335	2.358	1.557	3.331	2.371	1.571	3.317	2.335	1.518	3.290	2.349	1.531
H51 ATP	3.734	1.626	2.286	3.733	1.656	2.316	3.699	1.592	2.246	3.663	1.590	2.245
H52 ATP	2.827	1.779	0.760	2.825	1.776	0.786	2.823	1.767	0.705	2.775	1.805	0.715
O5' ATP	2.409	3.089	2.298	2.400	3.105	2.301	2.378	3.057	2.250	2.380	3.093	2.273
PA ATP	1.226	4.135	1.528	1.210	4.135	1.529	1.217	4.136	1.511	1.220	4.162	1.482
OA1 ATP	2.021	5.333	0.991	2.008	5.321	0.973	2.011	5.310	0.949	2.061	5.369	1.015
OA2 ATP	0.139	4.439	2.546	0.132	4.447	2.548	0.153	4.426	2.546	0.141	4.490	2.502
OA3 ATP	0.686	3.185	0.303	0.628	3.152	0.338	0.616	3.205	0.279	0.623	3.243	0.290
PB ATP	-0.360	2.396	-0.822	-0.371	2.401	-0.842	-0.376	2.445	-0.864	-0.492	2.326	-0.734
OB1 ATP	-1.028	3.551	-1.581	-1.032	3.559	-1.591	-1.136	3.557	-1.595	-1.124	3.462	-1.577
OB2 ATP	0.471	1.351	-1.586	0.485	1.361	-1.577	0.437	1.387	-1.594	0.420	1.343	-1.533
OB3 ATP	-1.371	1.489	0.087	-1.465	1.461	-0.046	-1.431	1.505	0.035	-1.496	1.513	0.180
PG ATP	-2.977	0.996	0.654	-2.977	0.990	0.695	-2.876	1.003	0.771	-3.782	0.849	0.660
OG1 ATP	-3.989	1.180	-0.532	-4.173	1.165	-0.386	-4.039	1.305	-0.422	-4.217	1.364	-0.907
OG2 ATP	-2.630	-0.505	0.920	-2.624	-0.505	0.899	-2.606	-0.491	0.905	-3.141	-0.515	0.720
OG3 ATP	-3.231	1.908	1.887	-3.127	1.934	1.896	-3.125	1.901	1.967	-3.801	1.880	1.758

PM3/MM Stepwise mechanism	I			TS2			P		
C1 DHA	-7.736	-0.982	0.962	-7.702	-0.961	0.973	-7.757	-1.018	0.963
C2 DHA	-6.217	-0.881	0.950	-6.185	-0.837	0.985	-6.240	-0.900	0.873
H21 DHA	-5.789	-1.241	1.908	-5.784	-1.093	1.987	-5.756	-1.386	1.744
H22 DHA	-5.840	-1.547	0.148	-5.790	-1.592	0.273	-5.918	-1.450	-0.034
C3 DHA	-8.517	-0.211	-0.080	-8.489	-0.215	-0.082	-8.596	-0.208	-0.002
H31 DHA	-8.320	0.874	0.044	-8.314	0.874	0.034	-8.459	0.870	0.225
H32 DHA	-8.173	-0.500	-1.092	-8.137	-0.507	-1.089	-8.252	-0.377	-1.040
O4 DHA	-8.283	-1.718	1.782	-8.259	-1.700	1.787	-8.268	-1.781	1.785
O5 DHA	-5.825	0.429	0.717	-5.789	0.433	0.615	-5.881	0.428	0.795
H5 DHA	-2.816	1.587	-0.612	-2.492	1.626	-0.444	-2.160	1.672	-0.213
O6 DHA	-9.891	-0.495	-0.084	-9.859	-0.523	-0.089	-9.954	-0.563	-0.026
H6 DHA	-10.276	-0.040	0.656	-10.254	-0.071	0.647	-10.328	-0.311	0.812
C1' ATP	6.235	2.222	-0.224	6.243	2.231	-0.276	6.236	2.209	-0.217
H1' ATP	7.279	2.577	-0.090	7.283	2.605	-0.160	7.280	2.569	-0.088
C2' ATP	5.512	3.096	-1.274	5.484	3.104	-1.301	5.505	3.081	-1.261
H2' ATP	5.227	2.509	-2.176	5.168	2.525	-2.197	5.203	2.495	-2.158
O2' ATP	6.342	4.073	-1.848	6.290	4.090	-1.895	6.327	4.056	-1.847
HT2 ATP	6.586	4.682	-1.161	6.559	4.691	-1.210	6.603	4.652	-1.161
C3' ATP	4.257	3.627	-0.531	4.253	3.637	-0.518	4.263	3.625	-0.504
H3' ATP	3.343	3.156	-0.971	3.321	3.183	-0.936	3.338	3.174	-0.943
O3' ATP	4.139	5.023	-0.655	4.151	5.035	-0.631	4.171	5.023	-0.617
HT3 ATP	3.363	5.242	-0.112	3.382	5.261	-0.085	3.395	5.254	-0.084
C4' ATP	4.437	3.193	0.955	4.470	3.190	0.958	4.444	3.175	0.976
O4' ATP	5.645	2.437	1.060	5.671	2.421	1.020	5.649	2.417	1.070
H4' ATP	4.609	4.062	1.630	4.661	4.055	1.634	4.619	4.035	1.663
C5' ATP	3.240	2.345	1.476	3.292	2.346	1.526	3.249	2.325	1.495
H51 ATP	3.591	1.545	2.156	3.671	1.590	2.241	3.593	1.534	2.187
H52 ATP	2.720	1.845	0.634	2.776	1.795	0.715	2.735	1.813	0.657
O5' ATP	2.345	3.078	2.245	2.378	3.084	2.275	2.345	3.064	2.252
PA ATP	1.202	4.195	1.490	1.226	4.166	1.512	1.226	4.196	1.523
OA1 ATP	2.062	5.394	1.059	2.040	5.369	1.022	2.052	5.390	1.053
OA2 ATP	0.089	4.432	2.502	0.116	4.445	2.505	0.093	4.402	2.496
OA3 ATP	0.679	3.408	0.183	0.686	3.276	0.248	0.710	3.385	0.183
PB ATP	-0.260	2.454	-0.928	-0.352	2.481	-0.862	-0.201	2.566	-0.963
OB1 ATP	-1.069	3.556	-1.701	-1.014	3.618	-1.662	-1.044	3.626	-1.683
OB2 ATP	0.749	1.586	-1.700	0.471	1.402	-1.577	0.662	1.548	-1.666
OB3 ATP	-1.177	1.496	-0.091	-1.452	1.658	-0.008	-1.206	1.569	-0.051
PG ATP	-4.120	0.727	0.910	-4.057	0.748	0.785	-4.139	0.768	0.801
OG1 ATP	-3.778	1.406	-0.603	-3.635	1.352	-0.659	-3.857	1.426	-0.593
OG2 ATP	-3.372	-0.592	0.981	-3.336	-0.601	0.971	-3.373	-0.592	0.916
OG3 ATP	-4.015	1.759	2.006	-3.977	1.776	1.927	-3.988	1.730	2.008

PM6/MM Concerted mechanism	R			TS_c			P		
C1 DHA	-7.037	1.692	0.836	-6.738	1.630	0.844	-6.365	1.528	0.872
C2 DHA	-5.962	0.622	0.700	-5.690	0.516	0.738	-5.279	0.478	0.670
H21 DHA	-6.021	-0.021	1.606	-5.869	-0.089	1.657	-5.427	-0.317	1.431
H22 DHA	-6.105	0.005	-0.208	-5.915	-0.121	-0.138	-5.364	0.026	-0.339
C3 DHA	-8.299	1.527	-0.009	-8.048	1.487	0.054	-7.672	1.382	0.081
H31 DHA	-9.213	1.584	0.615	-8.933	1.558	0.717	-8.547	1.456	0.756
H32 DHA	-8.341	2.338	-0.773	-8.114	2.308	-0.697	-7.745	2.202	-0.671
O4 DHA	-6.964	2.641	1.587	-6.647	2.611	1.557	-6.286	2.478	1.623
O5 DHA	-4.657	1.229	0.702	-4.341	0.935	0.795	-4.010	1.102	0.839
H5 DHA	-4.098	0.802	-0.051	-3.795	1.199	-0.397	-3.387	1.478	-0.946
O6 DHA	-8.254	0.317	-0.758	-8.082	0.287	-0.713	-7.714	0.179	-0.681
H6 DHA	-8.763	-0.413	-0.271	-8.648	-0.408	-0.241	-8.316	-0.491	-0.221
C1' ATP	5.102	1.618	-0.240	5.089	1.619	-0.248	5.070	1.634	-0.265
H1' ATP	5.933	2.167	0.263	5.924	2.169	0.245	5.909	2.190	0.217
C2' ATP	4.790	2.221	-1.619	4.769	2.218	-1.627	4.737	2.222	-1.645
H2' ATP	5.005	1.543	-2.474	4.978	1.535	-2.480	4.929	1.530	-2.495
O2' ATP	5.537	3.391	-1.926	5.522	3.381	-1.942	5.497	3.376	-1.975
HT2 ATP	5.992	3.809	-1.137	5.977	3.802	-1.156	5.960	3.799	-1.194
C3' ATP	3.287	2.627	-1.538	3.268	2.631	-1.537	3.241	2.649	-1.544
H3' ATP	2.627	1.944	-2.121	2.604	1.952	-2.119	2.565	1.984	-2.130
O3' ATP	3.083	3.890	-2.135	3.069	3.898	-2.127	3.051	3.926	-2.118
HT3 ATP	3.849	4.506	-1.961	3.847	4.505	-1.967	3.849	4.514	-1.982
C4' ATP	2.940	2.645	-0.031	2.926	2.645	-0.029	2.903	2.651	-0.037
O4' ATP	3.946	1.799	0.612	3.937	1.802	0.610	3.927	1.823	0.602
H4' ATP	3.051	3.665	0.397	3.036	3.664	0.399	2.992	3.668	0.399
C5' ATP	1.609	1.988	0.347	1.599	1.983	0.358	1.588	1.958	0.341
H51 ATP	1.753	0.991	0.816	1.756	0.994	0.838	1.783	0.972	0.811
H52 ATP	0.950	1.837	-0.541	0.938	1.828	-0.528	0.928	1.796	-0.541
O5' ATP	0.972	2.743	1.368	0.957	2.744	1.373	0.926	2.692	1.362
PA ATP	0.048	4.068	0.919	0.028	4.061	0.914	-0.013	4.005	0.904
OA1 ATP	1.000	5.059	0.283	0.981	5.054	0.284	0.916	4.979	0.217
OA2 ATP	-0.735	4.337	2.154	-0.772	4.342	2.135	-0.789	4.324	2.128
OA3 ATP	-1.002	3.638	-0.283	-1.032	3.598	-0.277	-1.132	3.491	-0.195
PB ATP	-1.544	2.400	-1.245	-1.611	2.398	-1.254	-1.528	2.514	-1.435
OB1 ATP	-2.684	2.957	-2.052	-2.681	3.026	-2.102	-2.560	3.314	-2.197
OB2 ATP	-0.327	1.880	-2.000	-0.378	1.836	-1.942	-0.286	1.967	-2.105
OB3 ATP	-2.204	1.278	-0.296	-2.507	1.275	-0.442	-2.466	1.257	-0.897
PG ATP	-1.880	0.317	1.099	-2.272	0.294	1.117	-2.650	0.160	1.308
OG1 ATP	-2.806	-0.845	0.685	-2.979	-0.984	0.661	-3.307	-1.245	1.204
OG2 ATP	-0.396	-0.085	1.025	-0.758	-0.021	1.086	-1.267	0.075	0.644
OG3 ATP	-2.266	1.099	2.330	-2.598	1.203	2.269	-2.536	0.821	2.685

PM6/MM Stepwise mechanism	TS1			I		
C1 DHA	-7.585	-0.858	0.855	-7.602	-0.872	0.860
C2 DHA	-6.073	-0.702	0.815	-6.086	-0.786	0.875
H21 DHA	-5.714	-0.750	1.876	-5.743	-0.799	1.927
H22 DHA	-5.670	-1.625	0.315	-5.692	-1.670	0.318
C3 DHA	-8.397	-0.125	-0.180	-8.399	-0.130	-0.182
H31 DHA	-8.273	0.970	-0.062	-8.276	0.964	-0.048
H32 DHA	-8.076	-0.416	-1.207	-8.074	-0.426	-1.206
O4 DHA	-8.110	-1.593	1.682	-8.138	-1.602	1.685
O5 DHA	-5.700	0.495	0.073	-5.691	0.413	0.219
H5 DHA	-5.083	1.811	-1.128	-4.827	2.067	-1.350
O6 DHA	-9.764	-0.515	-0.104	-9.769	-0.511	-0.113
H6 DHA	-10.216	-0.049	0.647	-10.208	-0.049	0.649
C1' ATP	6.203	2.142	-0.304	6.212	2.141	-0.305
H1' ATP	7.252	2.526	-0.295	7.262	2.524	-0.303
C2' ATP	5.370	2.950	-1.332	5.376	2.942	-1.330
H2' ATP	4.973	2.335	-2.168	4.966	2.323	-2.154
O2' ATP	6.209	3.882	-1.988	6.211	3.873	-1.996
HT2 ATP	6.421	4.646	-1.387	6.435	4.627	-1.385
C3' ATP	4.245	3.618	-0.500	4.258	3.612	-0.495
H3' ATP	3.222	3.375	-0.886	3.235	3.362	-0.875
O3' ATP	4.456	5.024	-0.617	4.474	5.016	-0.629
HT3 ATP	3.764	5.481	-0.023	3.800	5.480	-0.028
C4' ATP	4.448	3.169	0.963	4.461	3.166	0.968
O4' ATP	5.693	2.407	1.023	5.712	2.413	1.024
H4' ATP	4.613	4.026	1.656	4.612	4.026	1.658
C5' ATP	3.340	2.239	1.493	3.356	2.234	1.503
H51 ATP	3.785	1.541	2.236	3.821	1.521	2.224
H52 ATP	2.860	1.655	0.686	2.878	1.661	0.687
O5' ATP	2.366	2.965	2.230	2.384	2.965	2.238
PA ATP	1.316	4.009	1.353	1.306	4.024	1.364
OA1 ATP	2.155	5.241	0.998	2.118	5.282	1.036
OA2 ATP	0.163	4.247	2.300	0.149	4.257	2.315
OA3 ATP	1.053	3.153	0.023	1.064	3.181	0.028
PB ATP	-0.364	2.294	-0.724	-0.394	2.279	-0.749
OB1 ATP	-0.971	3.440	-1.532	-1.001	3.425	-1.554
OB2 ATP	0.318	1.196	-1.562	0.288	1.182	-1.589
OB3 ATP	-1.119	1.735	0.465	-1.107	1.736	0.461
PG ATP	-3.861	0.885	0.420	-3.989	0.863	0.414
OG1 ATP	-3.934	1.495	-1.114	-4.024	1.504	-1.136
OG2 ATP	-3.101	-0.396	0.458	-3.139	-0.369	0.436
OG3 ATP	-3.978	1.901	1.510	-4.002	1.902	1.493

AM1d/MM Solvent-assisted	R-w			TS1-w			I-w		
C1 DHA	-7.493	0.924	-0.260	-7.358	0.812	-0.442	-7.560	0.937	-0.227
C2 DHA	-6.207	1.638	-0.543	-6.010	1.321	-0.916	-6.283	1.654	-0.577
H21 DHA	-6.449	2.690	-0.883	-6.192	2.385	-1.253	-6.545	2.725	-0.837
H22 DHA	-5.540	1.671	0.381	-5.302	1.340	-0.044	-5.567	1.629	0.300
C3 DHA	-8.370	0.325	-1.324	-8.359	0.323	-1.464	-8.412	0.307	-1.301
H31 DHA	-9.077	-0.437	-0.886	-9.126	-0.337	-0.971	-9.122	-0.452	-0.877
H32 DHA	-8.956	1.150	-1.828	-8.874	1.222	-1.901	-8.977	1.128	-1.831
O4 DHA	-7.945	0.910	0.944	-7.690	0.863	0.756	-7.993	0.874	0.986
O5 DHA	-5.524	1.032	-1.655	-5.527	0.583	-1.985	-5.724	1.056	-1.690
H5 DHA	-4.645	0.644	-1.353	-4.449	0.297	-1.924	-3.264	0.171	-2.273
O6 DHA	-7.612	-0.395	-2.305	-7.768	-0.462	-2.469	-7.659	-0.415	-2.288
H6 DHA	-7.284	0.246	-3.019	-7.350	0.144	-3.113	-7.081	0.251	-2.799
C1' ATP	5.246	1.569	0.126	5.323	1.575	0.135	5.241	1.568	0.133
H1' ATP	5.984	2.254	0.647	6.118	2.220	0.613	5.972	2.256	0.659
C2' ATP	5.162	1.933	-1.367	5.279	1.914	-1.367	5.161	1.934	-1.361
H2' ATP	4.801	1.098	-2.036	4.896	1.062	-1.992	4.800	1.100	-2.030
O2' ATP	6.467	2.227	-1.879	6.563	2.185	-1.860	6.465	2.228	-1.871
HT2 ATP	6.860	3.036	-1.432	6.915	2.975	-1.408	6.860	3.037	-1.423
C3' ATP	4.181	3.125	-1.389	4.293	3.101	-1.428	4.181	3.128	-1.383
H3' ATP	3.607	3.210	-2.355	3.645	3.056	-2.341	3.611	3.218	-2.351
O3' ATP	4.964	4.328	-1.218	5.052	4.294	-1.410	4.962	4.329	-1.203
HT3 ATP	4.388	5.115	-1.418	4.468	5.006	-1.711	4.392	5.118	-1.416
C4' ATP	3.289	2.925	-0.144	3.469	2.974	-0.131	3.286	2.921	-0.142
O4' ATP	3.938	1.891	0.684	4.070	1.961	0.703	3.929	1.884	0.685
H4' ATP	3.262	3.841	0.519	3.500	3.927	0.459	3.252	3.836	0.523
C5' ATP	1.880	2.461	-0.502	2.029	2.547	-0.392	1.884	2.449	-0.508
H51 ATP	1.789	1.350	-0.664	1.982	1.450	-0.630	1.812	1.341	-0.688
H52 ATP	1.491	3.003	-1.415	1.614	3.124	-1.264	1.491	3.015	-1.405
O5' ATP	1.029	2.721	0.653	1.286	2.750	0.795	1.040	2.670	0.660
PA ATP	0.018	4.050	0.687	0.091	4.132	0.763	0.030	3.999	0.708
OA1 ATP	0.863	5.262	0.341	0.905	5.406	0.350	0.832	5.221	0.322
OA2 ATP	-0.755	3.944	1.976	-0.786	3.989	2.044	-0.764	3.917	1.978
OA3 ATP	-0.994	3.845	-0.561	-0.726	3.638	-0.715	-0.862	3.679	-0.625
PB ATP	-1.696	2.528	-1.281	-2.022	2.523	-1.209	-1.851	2.542	-1.217
OB1 ATP	-2.401	3.267	-2.409	-2.636	3.444	-2.316	-2.350	3.255	-2.452
OB2 ATP	-0.615	1.531	-1.646	-1.076	1.228	-1.891	-0.864	1.345	-1.691
OB3 ATP	-2.799	1.833	-0.373	-2.905	1.974	0.026	-2.922	2.092	-0.203
PG ATP	-3.085	0.574	0.738	-3.158	0.290	1.194	-3.174	0.529	0.803
OG1 ATP	-3.662	-0.515	-0.198	-3.783	-0.811	0.258	-3.558	-0.638	-0.096
OG2 ATP	-1.743	0.236	1.379	-1.653	0.078	1.650	-1.729	0.370	1.283
OG3 ATP	-4.151	1.070	1.702	-4.099	1.045	2.184	-4.148	1.012	1.836
Ow HOH	-2.304	-0.540	-2.443	-3.276	-0.250	-2.351	-2.628	-0.510	-2.371
Hw HOH	-1.844	0.227	-1.984	-1.650	0.416	-1.929	-1.314	0.443	-1.846
Hw HOH	-2.921	-0.901	-1.755	-3.234	-1.061	-1.873	-2.796	-1.151	-1.611

AM1d/MM Solvent-assisted	TS2-w			P-w		
C1 DHA	-7.446	0.907	-0.203	-7.250	0.869	-0.152
C2 DHA	-6.076	1.505	-0.429	-5.903	1.562	-0.291
H21 DHA	-6.141	2.285	-1.242	-5.871	2.133	-1.263
H22 DHA	-5.700	1.974	0.528	-5.723	2.236	0.595
C3 DHA	-8.322	0.340	-1.292	-8.158	0.371	-1.257
H31 DHA	-9.043	-0.413	-0.866	-8.902	-0.370	-0.837
H32 DHA	-8.876	1.193	-1.779	-8.708	1.249	-1.706
O4 DHA	-7.955	0.882	0.980	-7.777	0.807	1.017
O5 DHA	-5.242	0.458	-0.773	-4.869	0.582	-0.320
H5 DHA	-3.332	0.650	-2.095	-3.434	0.449	-2.477
O6 DHA	-7.613	-0.394	-2.304	-7.498	-0.359	-2.302
H6 DHA	-7.279	0.243	-3.011	-7.299	0.253	-3.083
C1' ATP	5.235	1.569	0.134	5.236	1.570	0.134
H1' ATP	5.964	2.259	0.660	5.964	2.259	0.663
C2' ATP	5.156	1.935	-1.360	5.160	1.936	-1.360
H2' ATP	4.796	1.100	-2.028	4.799	1.101	-2.028
O2' ATP	6.461	2.228	-1.870	6.466	2.228	-1.868
HT2 ATP	6.854	3.040	-1.425	6.857	3.040	-1.422
C3' ATP	4.176	3.128	-1.382	4.181	3.129	-1.384
H3' ATP	3.607	3.219	-2.351	3.616	3.222	-2.355
O3' ATP	4.956	4.329	-1.199	4.960	4.331	-1.199
HT3 ATP	4.382	5.118	-1.403	4.380	5.119	-1.389
C4' ATP	3.276	2.919	-0.145	3.277	2.917	-0.151
O4' ATP	3.920	1.882	0.683	3.921	1.883	0.681
H4' ATP	3.238	3.833	0.520	3.231	3.830	0.515
C5' ATP	1.876	2.447	-0.523	1.879	2.440	-0.535
H51 ATP	1.803	1.337	-0.685	1.806	1.327	-0.677
H52 ATP	1.499	2.997	-1.437	1.512	2.974	-1.461
O5' ATP	1.007	2.695	0.618	1.003	2.715	0.593
PA ATP	0.031	4.057	0.638	0.034	4.078	0.584
OA1 ATP	0.907	5.267	0.383	0.932	5.283	0.373
OA2 ATP	-0.798	3.932	1.887	-0.837	3.956	1.802
OA3 ATP	-0.761	3.901	-0.774	-0.728	3.966	-0.860
PB ATP	-1.818	2.806	-1.394	-1.812	2.871	-1.451
OB1 ATP	-1.985	3.317	-2.820	-1.691	3.075	-2.968
OB2 ATP	-0.872	1.486	-1.547	-1.065	1.475	-1.176
OB3 ATP	-3.064	2.725	-0.572	-3.187	3.160	-0.923
PG ATP	-3.403	0.121	1.006	-3.665	0.122	0.936
OG1 ATP	-3.679	-1.123	0.217	-3.848	-1.305	0.379
OG2 ATP	-1.925	0.290	1.259	-2.155	0.357	1.041
OG3 ATP	-4.342	0.814	1.917	-4.432	0.708	2.105
Ow HOH	-2.873	-0.147	-2.196	-2.885	-0.320	-2.217
Hw HOH	-1.373	0.615	-1.629	-1.608	0.668	-1.455
Hw HOH	-3.255	-0.853	-1.587	-3.288	-0.818	-1.424

AM1d/MM Solvent-assisted	R-w	TS_c-w	P´-w
C1 DHA	-7.334 -0.667 0.597	-7.354 -0.687 0.608	-7.371 -0.688 0.639
C2 DHA	-5.922 -0.355 0.186	-5.910 -0.390 0.395	-5.905 -0.424 0.533
H21 DHA	-5.466 0.364 0.929	-5.567 0.013 1.406	-5.566 -0.149 1.580
H22 DHA	-5.328 -1.318 0.136	-5.350 -1.327 0.088	-5.390 -1.350 0.126
C3 DHA	-8.513 -0.126 -0.149	-8.479 -0.113 -0.183	-8.471 -0.104 -0.186
H31 DHA	-8.622 0.984 0.010	-8.560 1.002 -0.023	-8.551 1.011 -0.030
H32 DHA	-8.421 -0.361 -1.251	-8.365 -0.356 -1.279	-8.353 -0.351 -1.281
O4 DHA	-7.523 -1.415 1.625	-7.619 -1.483 1.587	-7.696 -1.516 1.575
O5 DHA	-5.903 0.239 -1.126	-5.563 0.680 -0.482	-5.482 0.700 -0.243
H5 DHA	-5.048 0.753 -1.296	-4.964 0.443 -1.809	-4.890 0.420 -2.092
O6 DHA	-9.725 -0.795 0.255	-9.718 -0.752 0.188	-9.725 -0.729 0.164
H6 DHA	-10.100 -0.377 1.075	-10.068 -0.367 1.037	-10.067 -0.369 1.027
C1' ATP	6.139 2.145 -0.346	6.155 2.143 -0.344	6.159 2.142 -0.344
H1' ATP	7.179 2.594 -0.328	7.199 2.584 -0.330	7.203 2.581 -0.329
C2' ATP	5.263 2.886 -1.377	5.281 2.888 -1.370	5.287 2.889 -1.371
H2' ATP	4.950 2.259 -2.263	4.957 2.264 -2.251	4.964 2.265 -2.252
O2' ATP	6.043 3.915 -2.008	6.063 3.924 -1.989	6.071 3.925 -1.988
HT2 ATP	6.066 4.750 -1.472	6.016 4.758 -1.451	6.017 4.762 -1.454
C3' ATP	4.052 3.432 -0.586	4.079 3.440 -0.577	4.084 3.442 -0.579
H3' ATP	3.064 3.048 -0.993	3.086 3.053 -0.976	3.090 3.056 -0.977
O3' ATP	4.069 4.865 -0.739	4.087 4.870 -0.757	4.096 4.872 -0.762
HT3 ATP	3.262 5.244 -0.267	3.279 5.251 -0.264	3.293 5.258 -0.265
C4' ATP	4.296 3.080 0.895	4.327 3.091 0.907	4.331 3.095 0.905
O4' ATP	5.599 2.419 0.977	5.629 2.426 0.982	5.631 2.425 0.982
H4' ATP	4.395 4.004 1.549	4.447 4.019 1.548	4.456 4.022 1.546
C5' ATP	3.280 2.126 1.526	3.298 2.144 1.543	3.298 2.152 1.540
H51 ATP	3.763 1.573 2.378	3.771 1.566 2.387	3.765 1.574 2.386
H52 ATP	2.772 1.409 0.821	2.793 1.454 0.811	2.792 1.463 0.807
O5' ATP	2.269 2.909 2.214	2.276 2.933 2.203	2.276 2.943 2.197
PA ATP	1.184 3.837 1.352	1.321 3.973 1.285	1.326 3.988 1.277
OA1 ATP	1.861 5.139 0.944	2.109 5.262 1.028	2.133 5.269 1.027
OA2 ATP	-0.031 3.966 2.236	0.081 4.206 2.123	0.090 4.232 2.115
OA3 ATP	1.093 2.831 0.084	1.294 3.245 -0.152	1.271 3.245 -0.142
PB ATP	-0.271 2.361 -0.755	-0.094 2.281 -0.585	-0.118 2.264 -0.582
OB1 ATP	-0.717 3.632 -1.455	-0.906 3.398 -1.247	-0.933 3.384 -1.230
OB2 ATP	0.224 1.157 -1.549	0.321 1.167 -1.569	0.365 1.188 -1.577
OB3 ATP	-1.144 1.900 0.492	-0.266 1.676 0.807	-0.357 1.625 0.780
PG ATP	-2.785 1.295 0.448	-3.701 1.115 0.244	-3.773 1.094 0.200
OG1 ATP	-3.287 1.693 -0.936	-3.526 1.737 -1.126	-3.499 1.683 -1.221
OG2 ATP	-2.476 -0.196 0.633	-3.097 -0.268 0.396	-3.084 -0.257 0.344
OG3 ATP	-3.407 1.987 1.657	-3.968 1.884 1.504	-3.933 1.940 1.441
Ow HOH	-4.106 0.360 -3.045	-4.269 0.267 -2.787	-4.257 0.212 -2.814
Hw HOH	-3.330 0.658 -2.489	-3.323 0.670 -2.586	-3.322 0.804 -2.408
Hw HOH	-4.384 -0.541 -2.757	-4.165 -0.709 -2.961	-4.164 -0.777 -2.948

PM3/MM Solvent-assisted	I1-w			TS2-1-w			I2-w		
C1 DHA	-7.783	-0.980	0.976	-7.784	-0.987	0.989	-7.799	-0.988	0.955
C2 DHA	-6.263	-0.877	0.944	-6.264	-0.890	0.956	-6.277	-0.908	0.895
H21 DHA	-5.818	-1.322	1.857	-5.818	-1.350	1.862	-5.816	-1.494	1.715
H22 DHA	-5.902	-1.466	0.077	-5.906	-1.464	0.079	-5.942	-1.349	-0.064
C3 DHA	-8.578	-0.194	-0.044	-8.582	-0.186	-0.017	-8.591	-0.190	-0.058
H31 DHA	-8.404	0.889	0.117	-8.419	0.895	0.172	-8.400	0.891	0.104
H32 DHA	-8.227	-0.444	-1.063	-8.224	-0.407	-1.040	-8.249	-0.442	-1.080
O4 DHA	-8.318	-1.729	1.792	-8.321	-1.741	1.800	-8.334	-1.722	1.784
O5 DHA	-5.876	0.450	0.824	-5.880	0.439	0.863	-5.882	0.418	0.985
H5 DHA	-2.867	1.593	-0.571	-2.955	1.664	-0.615	-2.947	1.542	-0.580
O6 DHA	-9.945	-0.508	-0.062	-9.945	-0.514	-0.049	-9.962	-0.487	-0.067
H6 DHA	-10.339	-0.109	0.704	-10.349	-0.124	0.717	-10.353	-0.050	0.680
C1' ATP	6.225	2.217	-0.175	6.229	2.218	-0.189	6.222	2.224	-0.189
H1' ATP	7.264	2.575	-0.014	7.267	2.579	-0.024	7.260	2.585	-0.029
C2' ATP	5.524	3.099	-1.235	5.530	3.107	-1.246	5.521	3.102	-1.255
H2' ATP	5.249	2.517	-2.142	5.258	2.530	-2.157	5.253	2.517	-2.162
O2' ATP	6.371	4.070	-1.793	6.376	4.084	-1.795	6.362	4.079	-1.811
HT2 ATP	6.581	4.695	-1.110	6.565	4.719	-1.114	6.570	4.706	-1.129
C3' ATP	4.259	3.638	-0.514	4.262	3.641	-0.526	4.250	3.637	-0.541
H3' ATP	3.348	3.178	-0.974	3.354	3.181	-0.991	3.345	3.167	-1.001
O3' ATP	4.152	5.036	-0.631	4.151	5.039	-0.639	4.133	5.033	-0.670
HT3 ATP	3.373	5.257	-0.093	3.373	5.260	-0.102	3.355	5.259	-0.135
C4' ATP	4.406	3.190	0.971	4.405	3.188	0.958	4.398	3.203	0.948
O4' ATP	5.603	2.420	1.096	5.602	2.416	1.082	5.597	2.436	1.080
H4' ATP	4.570	4.052	1.658	4.568	4.048	1.648	4.558	4.073	1.626
C5' ATP	3.186	2.351	1.453	3.183	2.348	1.430	3.186	2.360	1.437
H51 ATP	3.513	1.527	2.116	3.506	1.517	2.086	3.516	1.539	2.101
H52 ATP	2.670	1.886	0.589	2.668	1.893	0.559	2.668	1.891	0.575
O5' ATP	2.291	3.079	2.225	2.284	3.067	2.207	2.283	3.078	2.212
PA ATP	1.159	4.217	1.485	1.160	4.223	1.499	1.166	4.234	1.508
OA1 ATP	2.043	5.400	1.052	2.045	5.391	1.053	2.041	5.393	1.043
OA2 ATP	0.073	4.467	2.526	0.087	4.463	2.544	0.080	4.463	2.533
OA3 ATP	0.564	3.456	0.193	0.519	3.473	0.194	0.534	3.459	0.199
PB ATP	-0.256	2.464	-0.985	-0.282	2.451	-0.929	-0.220	2.413	-0.898
OB1 ATP	-1.082	3.544	-1.785	-1.160	3.549	-1.783	-1.111	3.532	-1.775
OB2 ATP	0.847	1.700	-1.732	0.802	1.707	-1.698	0.852	1.717	-1.704
OB3 ATP	-1.175	1.460	-0.200	-1.234	1.495	-0.177	-1.188	1.471	-0.177
PG ATP	-4.167	0.737	0.999	-4.168	0.732	0.980	-4.168	0.686	1.051
OG1 ATP	-3.826	1.399	-0.523	-3.885	1.383	-0.552	-3.858	1.220	-0.510
OG2 ATP	-3.427	-0.588	1.069	-3.415	-0.591	1.046	-3.408	-0.629	1.157
OG3 ATP	-4.043	1.768	2.092	-4.029	1.761	2.077	-4.010	1.770	2.093
Ow HOH	-3.751	3.480	-2.211	-3.484	3.507	-2.319	-3.624	3.594	-2.469
Hw HOH	-4.024	2.907	-1.482	-3.836	2.938	-1.646	-3.940	3.382	-1.613
Hw HOH	-2.777	3.514	-2.158	-2.239	3.471	-1.978	-2.084	3.388	-1.913

PM3/MM Solvent-assisted	TS2-2-w	P-w
C1 DHA	-7.784 -1.040 0.998	-7.805 -1.009 0.987
C2 DHA	-6.266 -0.905 1.056	-6.287 -0.909 1.027
H21 DHA	-5.860 -1.411 1.954	-5.871 -1.322 1.965
H22 DHA	-5.849 -1.420 0.167	-5.885 -1.519 0.188
C3 DHA	-8.524 -0.316 -0.105	-8.553 -0.222 -0.070
H31 DHA	-8.290 0.767 -0.055	-8.346 0.861 0.057
H32 DHA	-8.179 -0.688 -1.089	-8.200 -0.505 -1.079
O4 DHA	-8.366 -1.759 1.811	-8.391 -1.732 1.796
O5 DHA	-5.919 0.429 1.082	-5.924 0.408 0.951
H5 DHA	-3.995 2.202 -1.171	-3.343 2.180 -1.560
O6 DHA	-9.909 -0.548 -0.114	-9.934 -0.480 -0.100
H6 DHA	-10.279 -0.073 0.620	-10.309 -0.071 0.671
C1' ATP	6.239 2.212 -0.180	6.253 2.217 -0.201
H1' ATP	7.279 2.569 -0.023	7.295 2.576 -0.055
C2' ATP	5.536 3.081 -1.248	5.539 3.101 -1.251
H2' ATP	5.269 2.489 -2.152	5.269 2.520 -2.161
O2' ATP	6.374 4.056 -1.811	6.368 4.089 -1.807
HT2 ATP	6.582 4.685 -1.131	6.464 4.785 -1.168
C3' ATP	4.265 3.614 -0.534	4.274 3.621 -0.517
H3' ATP	3.362 3.134 -0.987	3.367 3.140 -0.961
O3' ATP	4.142 5.009 -0.677	4.145 5.016 -0.653
HT3 ATP	3.373 5.239 -0.132	3.388 5.243 -0.092
C4' ATP	4.423 3.196 0.960	4.450 3.191 0.970
O4' ATP	5.618 2.425 1.091	5.644 2.416 1.079
H4' ATP	4.593 4.073 1.626	4.633 4.060 1.644
C5' ATP	3.213 2.366 1.476	3.239 2.363 1.489
H51 ATP	3.546 1.569 2.168	3.571 1.553 2.164
H52 ATP	2.696 1.864 0.634	2.710 1.879 0.643
O5' ATP	2.309 3.106 2.229	2.346 3.101 2.258
PA ATP	1.203 4.261 1.505	1.220 4.245 1.553
OA1 ATP	2.076 5.430 1.060	2.069 5.412 1.064
OA2 ATP	0.085 4.466 2.502	0.109 4.469 2.547
OA3 ATP	0.636 3.542 0.131	0.637 3.487 0.209
PB ATP	-0.036 2.354 -0.874	-0.093 2.388 -0.818
OB1 ATP	-1.124 3.385 -1.716	-0.970 3.595 -1.706
OB2 ATP	1.024 1.813 -1.805	0.883 1.734 -1.750
OB3 ATP	-0.763 1.320 -0.067	-1.009 1.477 -0.078
PG ATP	-4.199 0.789 0.875	-4.163 0.695 0.915
OG1 ATP	-4.200 1.264 -0.667	-4.150 1.103 -0.574
OG2 ATP	-3.426 -0.541 1.018	-3.359 -0.630 1.135
OG3 ATP	-3.983 1.848 1.967	-4.009 1.754 2.053
Ow HOH	-3.813 3.129 -1.925	-3.739 2.934 -2.067
Hw HOH	-3.927 3.885 -1.373	-4.300 3.364 -1.431
Hw HOH	-2.057 3.126 -1.676	-1.894 3.412 -1.704

7.2. Study of the phosphoryl transfer reaction from ATP to Dha in DHAK from *E. coli*

A theoretical study of the phosphoryl transfer reaction from ATP to Dha catalysed by DhaK from E. coli.

Isabel Bordes, Raquel Castillo and Vicent Moliner

ASAP in *Journal of Physical Chemistry B*

DOI**:** 10.1021/acs.jpcb.7b04862

THE JOURNAL OF PHYSICAL CHEMISTRY B

Subscriber access provided by UNIVERSITAT JAUME I

Article

A Theoretical Study of the Phosphoryl Transfer Reaction From ATP to Dha Catalyzed by DhaK From *Escherichia Coli*

Isabel Bordes, Raquel Castillo, and Vicent Moliner

J. Phys. Chem. B, **Just Accepted Manuscript** • DOI: 10.1021/acs.jpcb.7b04862 • Publication Date (Web): 29 Aug 2017

A Theoretical Study of the Phosphoryl Transfer Reaction from ATP to Dha Catalyzed by DhaK from *Escherichia coli*

I. Bordes,[1] R. Castillo[1,*] and V. Moliner[1,2,*]

1. Departament de Química Física i Analítica, Universitat Jaume I, 12071Castellón, Spain.

2. Department of Chemistry, University of Bath, Bath BA2 7AY, United Kingdom.

* to whom correspondence should be addressed:

R. Castillo: rcastill@uji.es

V. Moliner: moliner@uji.es

ORCID:

R. Castillo: 0000-0003-1597-1136

V. Moliner: 0000-0002-3665-3391

ABSTRACT

Protein kinases, representing one of the largest protein family involved in almost all aspects of cell life, have become one of the most important targets for the development of new drugs to be used in, for instance, cancer treatments. In this paper an exhaustive theoretical study of the phosphoryl transfer reaction from adenosine triphosphate (ATP) to dihydroxyacetone (Dha) catalyzed by DhaK from *Escherichia coli (E. coli)* is reported. Two different mechanisms, previously proposed for the phosphoryl transfer from ATP to the hydroxyl side chain of specific serine, threonine or tyrosine residues, have been explored based on the generation of free energy surfaces (FES) computed with hybrid QM/MM potentials. The results suggest that the *substrate-assisted* phosphoryl and proton-transfer mechanism is kinetically more favorable than the mechanism where an aspartate would be activating the Dha. Although the details of the mechanisms appear to be dramatically dependent on the level of theory employed in the calculations (PM3/MM, B3LYP:PM3/MM or B3LYP/MM), the transition states (TSs) for the phosphoryl transfer step appear to be described as a concerted step with different degrees of synchronicity in the breaking and forming bonds process in both explored mechanisms. Residues of the active site belonging to different subunits of the protein such as Gly78B, Thr79A, Ser80A, Arg178B and one Mg^{2+} cation would be stabilizing the transferred phosphate in the TS. Asp109A would have a structural role by posing the Dha and other residues of the active site in the proper orientation. The information derived from our calculations not only reveals the role of the enzyme and the particular residues of its active site, but it can assist in the rationally design of new more specific inhibitors.

INTRODUCTION

The transfer of a phosphoryl group from an ester or anhydride to a nucleophile[1] is involved in many processes taking place in the cells of living organisms, playing and essential role in propagation and signal transduction.[2-5] Protein kinases catalyze the transfer of phosphate groups to different substrates in cells, being the adenosine triphosphate (ATP) the most important biological phosphoryl donor. An indication of the relevance of kinases is the fact that there are more than 500 protein kinases encoded in the human genome. In addition, it has been found around 540 kinases in mice and 122 in yeast, representing one of the largest protein family.[6] Numerous kinases are closely involved in cancer progression.[7,8] The common feature maintained throughout protein kinase family is the sharing of a highly conserved catalytic center, making possible that one inhibitor molecule could bind to several kinases, blocking regular signaling needed for normal cellular function. However, it is also feasible that the capability of interfering in a multitude of kinases could be beneficial in cancer treatments, provided that the inhibitor is specific to cancerous cells.[2,3] During the last years many studies have been focused on drug research related with kinases and several small-molecule kinase inhibitor drugs have been approved for the treatment of different cancers.[9,10] Even so, despite its relevance for cellular control and their crucial contribution in modern medical techniques, the entire protein phosphorylation network has not been investigated deeply.[8]

Previous theoretical studies on phosphoryl transfer reactions have stressed the presence of low-lying d-orbitals on the phosphorus atoms that allows the existence of phosphorus pentavalent species as intermediates.[11-14] The reaction of phosphate transfers has been considered to occur through two alternative mechanisms;[15,16] associative, in which nucleophilic attack precedes the living group departure, and dissociative, in which departure of the living group takes place before the nucleophilic attack is produced. The reaction can in turn proceed through only one step but it does not necessarily imply a synchronous forming and breaking bonds process since associative-like or dissociative-like mechanism can be observed for concerted paths. Different computational studies of ATP hydrolysis indicate the dissociative mechanism energetically more favorable than the associative one.[17-19] Others show very small energetic difference between the associative, dissociative and concerted mechanisms.[20] And more recently, different

QM/MM studies[21-22] of the ATP hydrolysis in myosin indicates the flexibility of the hydrolysis mechanism of this enzyme which presented similar free energies, and the dissociative mechanism like the most favorable within a metaphosphate compound.

In addition, studies of guanosine triphosphate (GTP) hydrolysis catalyzed by different GTPases concluded that different types of GTP hydrolysis reaction pathways can be obtained in each system which indicates that enzymes catalyzing GTP hydrolysis are able to stabilize different transition states modulating the reaction path.[23,24]

Studies of the phosphate transfer from ATP to serine or tyrosine residues in protein kinases have been also widely reported.[3,25-31] Thus, a theoretical study in cyclin-dependent kinase (CDK2) by De Vivo et al.[25] proposed that an Asp residue located in the active site would play a structural role whereas the activation of the Ser nucleophile was proposed to be due to a proton transfer to ATP. The authors suggested a *substrate-assisted mechanism* through a single concerted transition state (TS). Smith and co-workers studied the phosphoryl transfer reaction to a Ser residue catalyzed by a CDK2 using QM/MM methods.[3] Their results support a concerted dissociative mechanism through a metaphosphate-like TS, where an Asp residue, rather than ATP as previously proposed, would be acting as the general base to activate the Ser nucleophile. They associated the conclusions of De Vivo et al. to the exclusion of the Asp from the QM region. Based also on QM/MM methods and classical MD simulations, a similar catalytic mechanism was suggested by Ojeda-May et al. in insulin receptor kinase (IRK).[26] In this case, a catalytic Asp residue would abstract the proton from a Tyr residue of the substrate peptide. Diaz and Field suggested the same catalytic role for Asp residue from DFT calculations on a cluster model of the active site of cAMP-dependent protein kinase (PKA), where the ATP phosphate would be transferred to a Ser side chain residue.[30] Later, Cheng and co-workers conducted combined *ab initio* QM/MM methods in the same enzyme supported likewise the role of the Asp residue as a catalytic base and describing the reaction mechanism as mainly dissociative.[31] Even so, the previously performed work of Hart et al.[27] supported the *substrate-assisted* reaction as well, since they obtained higher potential energy barriers when the proton was transferred from Ser to Asp in PKA. More recently, based on the generation of QM/MM PESs, Pérez-Gallegos et al. [28,29] analyzed the phosphate transfer from ATP to a Ser residue catalyzed by PKA and concluded that the catalytic mechanism was not substrate dependent, with the active site Asp acting as a general acid-base catalyst.

Another important debate on phosphate transfer reactions of ATP is the role of the two magnesium ions that appear in the active site of kinases. It has been proposed that they are useful not only to stabilize ATP binding but also to accelerate the chemical reaction.[32,33] These metal ions have been reported to be greatly important to stabilize the negative charges accumulated during the phosphate transfer and before.[34,3] Its indispensability in many enzymatic systems has motivated several authors to determine the specific roles of Mg^{2+} in catalysis. Nevertheless, as stated by Lopata et al., despite the indispensability of the metal ion in most phosphoryl transfer and hydrolysis catalyzing enzymes, its role was not completely clear.[35]

Dihydroxyacetone kinases (DhaK) belong to a family of sequence-conserved enzymes that phosphorylate dihydroxyacetone (Dha) converting it in dihydroxyacetone phosphate (Dha-P).[34,36-38] Dha-P is an intermediate for the synthesis of pyruvate[34] and it is a very valuable compound in nature since it is used as phosphoryl donor in several enzyme-catalyzed aldol reactions for the formation of C-C bonds.[39,40] DhaK from *Escherichia coli* (*E. coli*), the most characteristic of the PTS-dependent DhaKs, is constituted by DhaK, DhaL and DhaM subunits.[41] The reaction starts with phosphorylation of DhaM, by the small phospho-carrier protein HPr of the PTS. Phosphorylated DhaM forms a complex with the ADP-bound DhaL subunit and transfers the phosphate to ADP turning it into ATP. Then, the ATP linked DhaL associates with DhaK subunit, which enclose the Dha compound bounded through a hemiaminal bond to His218. Finally, phosphate is transferred from ATP to Dha yielding as the product of the reaction, Dha-P and ADP.[34,38,41,42]

The molecular mechanism of the enzyme catalyzed phosphate transfer from ATP to Dha has been proposed based on the crystal structure of the *E. coli* DhaK-DhaL complex bound to ADP and Dha determined by Shi and co-workers.[34] The analysis of experimental results allows them to present a general catalytic mechanism for DhaK in which residues His56, His218 and Asp109 would be significantly involved in the reaction progress. In particular, they divided the whole reaction in three stages: 1) binding of Dha to enzyme through a hemiaminal bond with His218, acting the His56 as an acid donating its proton to Dha substrate; 2) Phosphate transfer from ATP to Dha where Asp109 acts as a base in order to prepare the Dha for the attack; and 3) release of the reaction product, Dha-P, from the enzyme.[34]

Nevertheless, and despite the great contributions to the study of kinases, the explicit reaction between ATP and Dha in DhaK from *E.coli* has not been studied by means of

computational methods. The analysis of the reaction by computational tools in DhaK can provide a better understanding of the behavior of these enzymes and a valuable contribution to clarify the molecular mechanism of phosphoryl transfer reactions. The recent QM/MM computational study of the counterpart reaction in aqueous solution carried out in our laboratory can be used to establish the methodological basis for the study of the reaction in the active site of DhaK and to estimate its catalytic efficiency.[43] In the present paper, we present an exhaustive theoretical study of the phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*. The possible reaction mechanisms are described based on the exploration of Free Energy Surfaces (FESs) computed by means of MD simulations with hybrid QM/MM potentials. The information derived from the calculations will allow not only revealing the role of the enzyme and the particular residues of the active site, but can have potential applications in the rational design of new more specific inhibitors.

COMPUTATIONAL METHODS

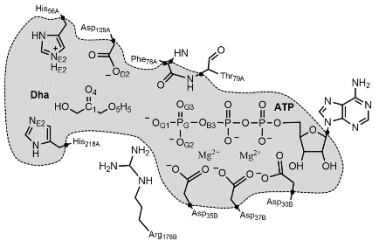
The initial coordinates were taken from the X-ray structure of *E. coli* DhaK, with PDB entry 3PNL.[34] This structure has two domains, DhaK and DhaL, differentiated as chain A and chain B. The chain A has 356 amino acids and the glycerol substrate while chain B contains 211 amino acids, ADP and two magnesium ions. The original ADP molecule was modified to ATP by adding a phosphate group manually within the help of Molden program[44]. The glycerol molecule was modified to Dha changing one alcohol functional group by a ketone group.

The protonation state of titratable residues has been determined using the PropKa program of Jensen *et al.*[45,46] at pH = 7. According to the results, all residues were found at their standard protonation state in aqueous solution, except His56A that had to be double protonated. The coordinates of the hydrogen atoms were added using the fDYNAMO library.[47] Then, 21 sodium ions were added to neutralize the negative charges of the system. Afterwards, the system was solvated with a pre-equilibrated orthorhombic box of water molecules with dimensions 100x80x80 $Å^3$. Water molecules with an oxygen atom lying within 2.8 Å of any heavy atom of ATP, Dha or protein were removed. The resulting system consists in 62823 atoms.

Once the model was generated, in order to relax energetically unfavorable interactions, series of optimizations using the conjugate gradient method followed by 5 ns of

classical MD simulations in the NVT ensemble at a temperature of 300 K using the Langevin-Verlet algorithm using a time step of 1 fs. The OPLS-AA force field,[48] and the TIP3P[49] force fields, as implemented in NAMD parallel molecular dynamics code,[50] were used to describe the protein and the water molecules, respectively. Parameters for Dha were generated with the ParamChem application[51,52] while parameters from the CHARMM general force field v. 2b7[53] were used for the ATP. Cut-offs for nonbonding interactions between atoms were applied using a switching-force scheme, within a range radius from 12 to 15 Å with periodic boundary conditions. The time evolution of the RMSD of the backbone atoms of the enzyme as well as the RMSD computed for just the backbone atoms of the active site residues (those in a radios sphere of 8 Å) during 5 ns of MD simulations show that the system can be considerate as equilibrated after the MD simulations (see Figures S1 and S2 of Supporting Information). This conclusion is supported by the inspection of the time evolution of individual key interactions established in the active site during the equilibration MD simulation (see Figures S3, S4 and S5 of Supporting Information). The resulting system was taken as the initial structure for the following QM/MM calculations. The Dha molecule, part of ATP (phosphate groups and ribose ring), the side chains of residues involved in the reaction mechanism (His218A, His56A and Asp109A), the magnesium ions and part of the residues in the coordination sphere of these magnesium ions (Asp30B, Asp35B, Asp37B, Phe78A and Thr79A) were treated quantum mechanically (see Scheme 1). The rest of the atoms of the protein and the water molecules were treated by means of the OPLS-AA and TIP3P[49] force field, respectively. To saturate the valence of the QM/MM frontier we used the link atom procedure[54], placing this atom between C of the ribose molecule and N of the adenine molecule in the case of ATP, between C_α and C_β in aspartate and histidine residues, between carboxyl carbon and C_α in Phe78A and between amino nitrogen and C_α in Thr79A (see Scheme 1). Thus, the QM part involves 96 atoms with a total charge of -3.

For all simulations, any residues further than 25 Å from the O_{B3} oxygen atom of ATP were kept frozen during the simulations. As well as in NAMD calculations, cut-offs for nonbonding interactions were applied using a switching function, within a radius range from 12.0 to 15.0 Å, employing periodic boundary conditions.



Scheme 1. Schematic representation of the active site of the DhaK. Grey region contains atoms treated quantum mechanically. Link atoms are represented as black dots.

The OPLS-AA and TIP3P force fields, as implemented in fDYNAMO library, were used to describe the protein atoms of the MM region and the water molecules, respectively. The semiempirical Parameterized Method Number 3 (PM3)[55], as implemented in fDYNAMO, was initially selected to treat the QM region during the QM/MM calculations. PM3 is typically 3-4 orders of magnitude faster than DFT methods and has been shown to produce a huge stabilization in the energy of phosphorane compounds, resulting from the use of a minimal valence basis.[56] Therefore, it has been used extensively to successfully model phosphorus and phosphate groups.[57-60] In fact, the semiempirical Hamiltonian PM3 was already tested to study the counterpart reaction in solution in our group, by means of QM/MM models,[43] showing that optimized structures with PM3/MM method rendered structures with all P-O bonds with reasonable interatomic distances. It is important to point out that, despite AM1d/MM calculations provided more accurate energy values for the study of the reaction in solution,[43] considering that the energies obtained at low-level (LL)/MM are later improved at high-level (HL)/MM with spline corrections (as explained below), we selected the PM3 Hamiltonian as the LL QM method due to the more reasonable structures obtained in the protein environment when combining with the MM force fields during the QM/MM MD simulations and energy optimizations.

In order to explore the full reaction mechanism of the DhaK catalysed reaction, PESs were first performed by scanning the appropriate combination of the interatomic distances to explore every single possible chemical step. Later FESs were computed in terms of potentials of mean force (PMFs). The umbrella sampling approach[61] was used

to constrain the system along the selected values of the reaction coordinates by employing a force constant of 2500 kJ mol^{-1} $Å^{-2}$. The value of the force constant used for the harmonic umbrella sampling were determined to allow full overlap of the different windows traced in the PMF evaluation, but without losing control over the selected coordinate. Each window consisted of 15 ps of equilibration followed by 20 ps of production. The Verlet algorithm was used to update the velocities. The initial structures in each window of the 2D-PMFs were selected from the corresponding previously generated PESs at the corresponding level of theory, whereas the 1D-PMFs were performed starting from the transition state structure. The probability distributions were put together by means of the weighted histogram analysis method (WHAM),[62] to obtain the full probability distribution along the reaction coordinate.

The generation of FESs requires the evaluation of a large number of structures by means of QM/MM MD simulations. Consequently, calculations are usually restricted to the use of semiempirical Hamiltonians, as it is the case in the present study. In order to reduce the errors associated with the quantum low-level PM3 employed in our simulations, following the studies of Truhlar et al.,[63-65] an interpolated correction term was applied to any value of the reaction coordinate ξ, selected to generate the FES. Then, a continuous new energy function was generated that corrects the PMFs, as previously performed in our laboratory:[66-68]

$$E = E_{LL/MM} + S[\Delta E_{LL}^{HL}(\xi)] \quad (1)$$

where S denotes a spline under tension function,[69,70] and its argument is a correction term evaluated from the single-point energy difference between a high-level (HL) and a low-level (LL) calculation of the QM subsystem. The semiempirical PM3 Hamiltonian was used as LL method, while the B3LYP method was selected for the HL energy calculation employing the 6-31G(d,p) basis set. S is adjusted to a defined grid depending on the reaction step studied. HL single energy calculations are computed on optimized geometries obtained in the corresponding PESs at LL.

In those chemical steps explored by means of 2D-PMFs, the correction term is expressed as a function of two coordinates ξ_1 and ξ_2, being the new energy function:

$$E = E_{LL/MM} + S[\Delta E_{LL}^{HL}(\xi_1, \xi_2)] \quad (2)$$

In equation 2, S is adjusted to a set of points corresponding to the HL single energy calculations on geometries optimized at PM3 method, as previously explained. All the

B3LYP calculations were carried by combining fDYNAMO with Gaussian09 program.[71]

RESULTS AND DISCUSSION

As stated in the Introduction section, the aim of the work presented here is the study of the molecular mechanism of the phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK. A detail of the active site of the starting structure in the reactant state, after equilibrated through QM/MM MD simulations as described in previous section, is shown in Figure 1. As observed in the figure, the reacting species are well oriented for the phosphoryl transfer. ATP is coordinated to the Mg^{2-} cations while Dha is stabilized by hydrogen bond interactions, mainly with Asp109A and His56A. The stability of these interactions is confirmed not only by the analysis of this snapshot but by the time evolution of these distances, as deposited in the Supporting Information. This initial structure was used as starting point to generate PESs for every chemical step involved in the mechanisms to be studied (PM3/MM PESs are deposited as Figures S3, S4 and S5 in the Supporting Information material).

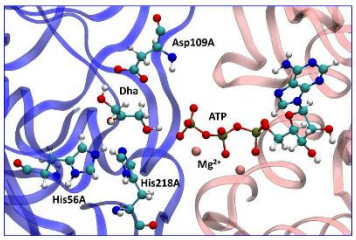
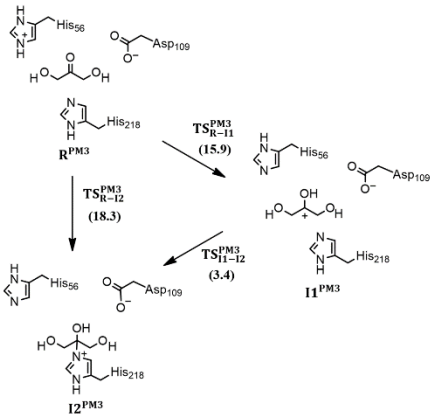


Figure 1. Representative snapshot of the active site of the DhaK in the reactants state after the QM/MM MD simulations. ATP, Dha, two Mg^{2+} ions and the key amino acids involved in the reaction mechanisms are shown as balls and sticks.

In the initial catalytic step, as depicted in Scheme 2, Dha is anchored to the enzyme through the formation of a covalent bond between the nitrogen atom NE2 of His218A and the carbon atom C1 of Dha. The nearby His56A acts as an acid, transferring the proton $HE2_{His56}$ to the oxygen atom O4 of Dha. The distance between nitrogen atom of His218A and the carbon atom of Dha, d($NE2_{His218}$-C1), and the antisymmetric

combination of the distances describing the hydrogen transfer from His56A to Dha, d($NE2_{IIis56}$-$HE2_{IIsp56}$)-d(O4-$HE2_{IIis56}$), were used as distinguished reaction coordinates to generate a 2D-PMF. The resulting FESs at PM3/MM level and with spline corrections at B3LYP:PM3/MM level are shown in Figure 2a and 2b, respectively, while the free energy barriers are reported in Scheme 2.



Scheme 2. Schematic representation of the initial step of the phosphorylation reaction mechanism corresponding to the formation of a covalent bond between Dha and the enzyme DhaK from *E. coli*. Values of the free energies barriers computed at PM3/MM level, in kcal·mol^{-1}, are reported in brackets.

The analysis of Figure 2a reveals that the transformation from reactants to the intermediate $I2^{PM3}$ can take place through a stepwise or a concerted mechanism. In the former, first the proton from His56A is transferred to Dha, obtaining a stable intermediate $I1^{PM3}$. Then the formation of a covalent bond between Dha and His218A takes place through the TS^{PM3}_{I1-I2} reaching the intermediate $I2^{PM3}$. The proton transfer can also take place concomitant with the $NE2_{His218}$-C1 covalent bond formation through the concerted TS^{PM3}_{R-I2}. Analysis of the structure of TS^{PM3}_{R-I2} shows that the bond formation between Dha and His218A is in a very advanced stage of the reaction (distance $NE2_{His218}$-$C1_{Dha}$ is 1.58 Å) while the transferring proton is located between its donor and

acceptor atom (distances $NE2_{His56}$-$HE2_{His56}$ and $O4_{Dha}$-$HE2_{His56}$ are 1.22 and 1.29 Å, respectively). A full list of the key interatomic distances of the located transition states on the FES are reported in Table S1 of the Supporting Information. From the energetic point of view, the free energy barrier of the rate limiting step of the stepwise mechanism, obtained at PM3/MM level, is only slightly lower than the concerted mechanism (15.9 and 18.3 kcal·mol^{-1}, respectively) suggesting the possibility of competitive mechanisms. Nevertheless, the corrected FES at B3LYP:PM3/MM level (Figure 2b) shows the absence of the intermediate $I1^{PM3}$, and the emergence of two concerted but very asynchronous mechanisms. Interestingly, the free energy barriers, 11.8 and 10.8 kcal·mol^{-1}, are significantly lower than the values obtained at PM3/MM level. The NE2-C1 bond is almost formed in the $TS1_{R-I2}^{DFT:PM3}$ but the proton transfer from His56A to Dha is in an early stage of the process. By contrast, $TS2_{R-I2}^{DFT:PM3}$ describes a situation where the proton transfer is almost completed but the formation of the covalent bond between Dha and His218A is in an early stage of the process.

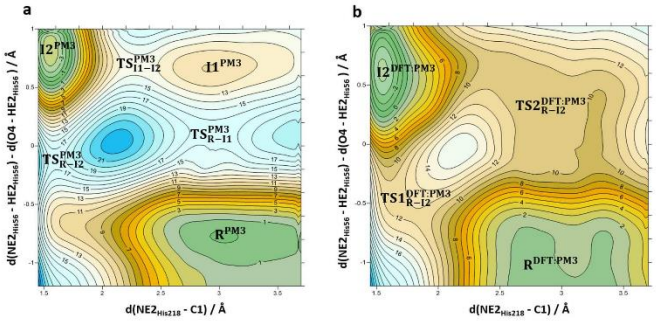


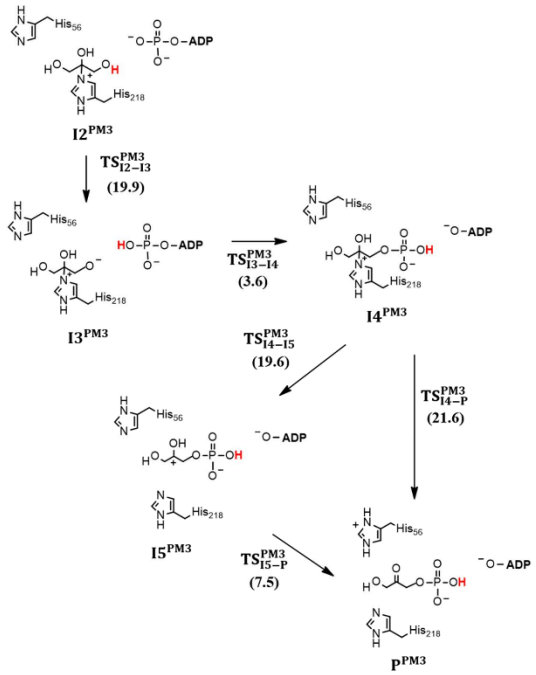
Figure 2. FESs of the first step of the phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli* corresponding to the formation of a covalent bond between Dha and the enzyme. Surfaces obtained as 2D-PMFs at PM3/MM level (panel a) and with spline corrections at B3LYP:PM3/MM level (panel b). Values on the isoenergetic lines are in kcal·mol^{-1}.

Then, after describing the first step of the DhaK in which the Dha is bounded to His218A (see Scheme 2), the rest of the reaction mechanism has been studied at three

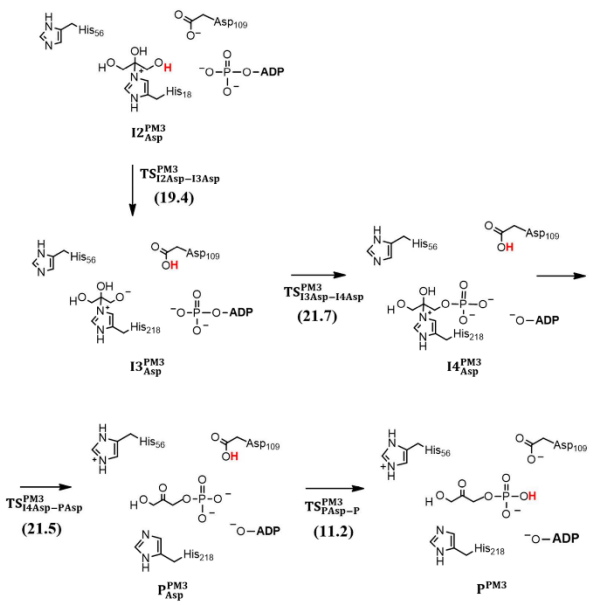
levels of theory, exploration of the FESs at PM3/MM, correction of the resulting FESs at DFT:PM3/MM level, and computing of PESs at DFT/MM level.

PM3/MM FESs of the phosphorylation reaction.

Once Dha is bounded to the enzyme ($I2^{PM3}$), the γ-OH group of Dha should be activated in order to make feasible the phosphate transfer from ATP. Thus, in the *substrate-assisted mechanism* the proton from Dha is directly transferred to the ATP whereas Asp109A has just a structural role (see Scheme 3). In the *asp-assisted mechanism,* Asp109A abstracts the proton from Dha and generates an oxyanion suitable for the attack of the ATP γ-phosphorus atom (see Scheme 4). As mentioned in the Introduction section, the first mechanism was supported by theoretical studies of the phosphoryl transfer reaction from ATP to the hydroxyl side chain of specific serine, threonine or tyrosine residues, while the later was proposed by Shi and co-workers based on crystal structures analysis.[34]



Scheme 3. Schematic representation of the *substrate-assisted mechanism* for the phosphorylation reaction from $I2^{PM3}$ intermediate to products, P^{PM3} in DhaK from *E. coli* obtained at PM3/MM level. Values of free energy barriers, in kcal·mol^{-1}, are reported in brackets.



Scheme 4. Schematic representation of the *asp-assisted mechanism* for the phosphorylation reaction from $I2^{PM3}_{Asp}$ intermediate to products, P^{PM3} in DhaK from *E. coli* obtained at PM3/MM level. Values of the free energy barriers, in kcal·mol^{-1}, are reported in brackets.

Substrate-assisted mechanism. A 2D-PMF was generated to explore the proton transfer from Dha to ATP and the phosphate transfer from ATP to Dha. Thus, starting from the $I2^{PM3}$ structure, two anti-symmetric combinations of distances, d($O5_{Dha}$-$H5_{Dha}$)-d($OG3_{ATP}$-$H5_{Dha}$) and d(PG_{ATP}-$OB3_{ATP}$)-d(PG_{ATP}-$O5_{Dha}$), were used to generate the FES. In the next stage, consisting into the release of the phosphorylated Dha from the protein, the anti-symmetric combination of distances defining the proton transfer back to the His56A, d($O4_{Dha}$-$HE2_{His56}$)-d($NE2_{His56}$- $HE2_{His56}$), together with the inter atomic

distance defining the $C1_{Dha}$ - $HE2_{His56}$ breaking bond, were used to generate the FES. The resulting FESs computed at PM3/MM level are shown in Figure 3.

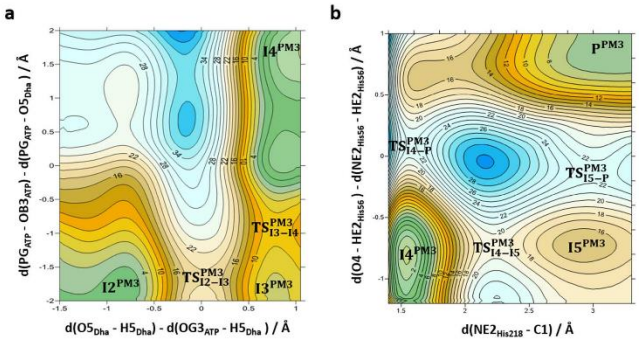


Figure 3. PM3/MM FESs of the *substrate-assisted mechanism* for the phosphorylation reaction catalyzed by DhaK from *E.coli*. a) Proton transfer from Dha to the transferring phosphoryl group of ATP followed by the phosphoryl transfer to Dha; and b) releasing of the final products where the proton returns to His56A and the bond with His218A is cleft. Values of the isoenergetic lines on the 2D-PMFs are in kcal·mol.$^{-1}$

The first conclusion that can be derived from the analysis of the FESs presented in Figure 3 is that, according to panel a, the activation of the phosphate by the transferring of a proton from Dha to the phosphate of ATP ($I2^{PM3}$ to $I3^{PM3}$) does not take place concertedly with the phosphoryl transfer itself ($I3^{PM3}$ to $I4^{PM3}$). Interestingly, the phosphate transfer step presents free energy barriers significantly lower than the rest of the steps. Indeed, while the rest of the steps have free energy barriers in the range between 19.6 to 21.6 kcal·mol^{-1}, the free energy barrier of the phosphoryl transfer step from the intermediate $I3^{PM3}$ is just 3.6 kcal·mol^{-1}. The complete migration of the phosphate group is reached in the $I4^{PM3}$ intermediate that appears to be much more stable than the $I3^{PM3}$ intermediate and even more stable than the initial $I2^{PM3}$ intermediate. Key interatomic distances of the transition state structures located on the quadratic regions of the FESs are reported in Table S2 of the Supporting Information. A representative structure of the TS of the phosphoryl transfer step, $I3^{PM3}$ to $I4^{PM3}$, is presented in Figure 4. The TS^{PM3}_{I3-I4} describes an asynchronous breaking and forming

bonds process with the phosphoryl transfer in an early stage of the process. The distances from the transferring phosphorous atom to the donor and acceptor atoms are 1.77 ± 0.04 and 2.57 ± 0.05 Å, respectively, with a distance between the donor and acceptor oxygen atoms of 3.98 ± 0.09 Å, and an angle of $OB3_{ATP}$-PG_{ATP}-$O5_{Dha}$ of 134.8 degrees. Hydrogen bond interactions are established in this TS between the oxygen atoms of the phosphate and residues Ser80A, Thr79A and Arg178B, together with interactions with a Mg^{2-} cations and a C-H bond of Gly78B. The Asp109A, on the contrary, is interacting with the substrate. It should be mentioned that, based on activity measurements on site-directed mutants, Shi et al.[34] proved that the conserved residues His218A, His56A and Asp109A were important for the catalytic activity of the DhaK, which would be in agreement with the description of TS_{I3-I4}^{PM3}.

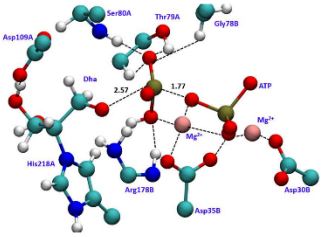


Figure 4. Representative snapshot of the TS_{I3-I4}^{PM3} structure obtained at PM3/MM corresponding to the phosphoryl transfer step, $I3^{PM3}$ to $I4^{PM3}$, in the *substrate-assisted mechanism*. Dashed lines represent important interactions and distances are reported in Å.

As observed in Figure 3b, the release of the products from the protein can be reached by means of a stepwise mechanism through an intermediate $I5^{PM3}$ (with free energy barriers of 19.6 and 7.5 kcal·mol^{-1}) or in a competitive single step (with a free energy barrier of 21.6 kcal·mol^{-1}).

Asp-assisted mechanism. In order to study the first step of this mechanism, and based on the experience achieved from the study of the previous mechanism, a 1D-PMF has been explored using the antisymmetric combination of distances corresponding to the transfer of proton from Dha to Asp109A, d($O5_{Dha}$-$H5_{Dha}$)-d($H5_{Dha}$-$OD2_{Asp109}$). Next, the phosphate is transferred from ATP to Dha which corresponds to the conversion from the

intermediate $I3^{PM3}_{Asp}$ to the intermediate $I4^{PM3}_{Asp}$ in Scheme 4. In this case, a 1D-PMF was performed using the antisymmetric combination d(PG_{ATP}-$OB3_{ATP}$)-d(PG_{ATP}-$O5_{Dha}$). Any attempt to search for a concerted proton and phosphate transfer process, by means of exploring 2D-FESs, was unsuccessful. After the phosphate transfer, Dha-P dissociates from the enzyme by breaking the hemiaminal bond with His218A and the abstraction of the proton by His56A. This can be considered as the inverse step of the initial step of the reaction mechanism shown in Scheme 2. Therefore, a 2D-PMF was explored to study the release of Dha-P from the enzyme and arriving to the products, P^{PM3}_{Asp}. The employed distinguished reaction coordinates are the distance d($NE2_{His218}$-C1) and the anti-symmetric combination d($NE2_{His56}$-$HE2_{His56}$)-d(O4- $HE2_{His56}$). At this point, we must bear in mind that Asp109A is still protonated whereas in the starting structure it was deprotonated. Consequently, a migration of this proton to one of the oxygen atoms of phosphate group of Dha-P ($OG2_{ATP}$) has been additionally studied, what is referred in Scheme 4 as the conversion from P^{PM3}_{Asp} to P^{PM3}. Hence, a 1D-PMF was computed using the anti-symmetric combination of distances d($OD2_{Asp109}$-$H5_{Dha}$) - d($H5_{Dha}$-$OG2_{ATP}$). The corresponding FESs for the full transformation from $I2^{PM3}_{Asp}$ to P^{PM3} computed at PM3/MM level are displayed in Figure 5 and the free energy barrier of every step are reported in Scheme 4.

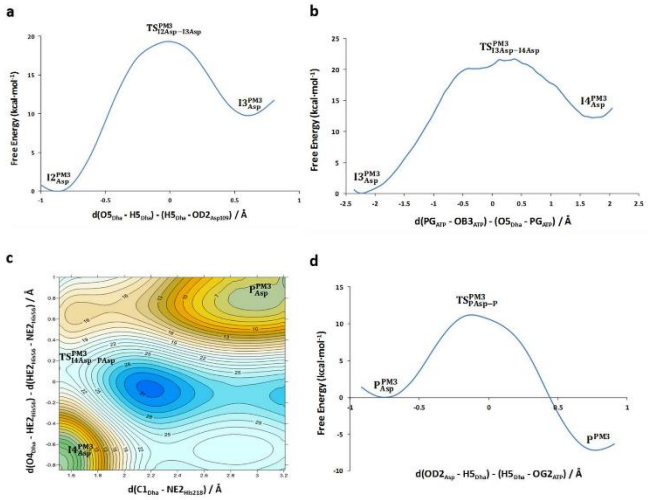


Figure 5. PM3/MM FESs of the chemical steps of the *asp-assisted mechanism* for the phosphorylation reaction catalyzed by DhaK from *E.coli*. Values of the isoenergetic lines on the 2D-PMF are in kcal·mol^{-1}. a) Proton transfer from Dha to Asp109A; b) phosphoryl transfer from ATP to Dha; c) releasing of the products where the proton returns to His56A and the cleavage of the bond with His218A is cleft; and d) proton transfer from Asp109A to ADP, reaching the final products.

The analysis of the FESs presented in Figure 5 shows how the phosphoryl transfer step, from $I3_{Asp}^{PM3}$ to $I4_{Asp}^{PM3}$, presents the highest energy barrier, 21.7 kcal·mol^{-1}. Nevertheless, the transformation from $I2_{Asp}^{PM3}$ to $I3_{Asp}^{PM3}$, and especially the transformation from $I4_{Asp}^{PM3}$ to P_{Asp}^{PM3} also take place through similar free energy barriers; 19.4 and 21.5 kcal·mol^{-1}, respectively. A detail of a representative structure of the TS for the phosphoryl transfer step is shown in Figure 6 while key interatomic distances of all transition states are reported in Table S3 of Supporting Information. As in the case of the *substrate-assisted mechanism*, the transferred phosphate is stabilized in the $TS_{I3Asp-I4Asp}^{PM3}$ by hydrogen bond interactions with three residues, Ser80A, Thr79A and Arg178B, and by one of the Mg^{2+} cations located in the active site. Gly78B also interacts with the transferring

phosphate, although by a weaker C-H···O interaction. The protonated Asp109A is interacting with Dha but, as in the previous mechanism, with the hydroxyl group. According to the distances established between the phosphor atom with the donor and acceptor oxygen atoms (2.32 ± 0.06 and 1.98 ± 0.06 Å, respectively), we would be describing this process as a concerted asynchronous phosphoryl transfer but, in contrast to the *substrate-assisted mechanism*, the TS would be in an advance stage of the reaction. The distance between the donor and acceptor oxygen atoms (4.07 ± 0.11 Å) and the $OB3_{ATP}$-PG_{ATP}-$O5_{Dha}$ angle on the $TS^{PM3}_{I3Asp-I4Asp}$ (158.7 degrees) are similar to the values obtained in the TS of the phosphoryl transfer step obtained in the *substrate-assisted mechanism*, within the statistical uncertainty.

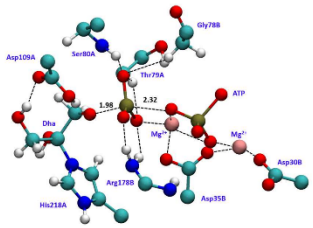


Figure 6. Representative snapshot of the $TS^{PM3}_{I3Asp-I4Asp}$ obtained at PM3/MM level corresponding to the phosphoryl transfer step, $I3^{PM3}_{Asp}$ to $I4^{PM3}_{Asp}$, in the *asp-assisted mechanism*. Dashed lines indicate important interactions and key distances are reported in Å.

In addition to the exploration of the two mechanisms, a reaction path from $I4^{PM3}_{Asp}$ to $I4^{PM3}$ has been traced by computing a 1D-PMF using the antisymmetric combination of the bond-breaking and bond-forming distances describing the proton transfer from Asp109A to one of the negatively charged phosphate oxygen atoms of Dha-P: d($OD2_{Asp109}$-$H5_{Dha}$)-d($H5_{Dha}$-$OG2_{ATP}$). The resulting free energy profile, presented in Figure 7, reveals that the $I4^{PM3}$ intermediate located when exploring the *substrate-assisted mechanism* is significantly more stable than the $I4^{PM3}_{Asp}$. This, together with the comparison of the energy barriers for each step, supports the proposal of the *substrate-assisted mechanism* as being more favourable than the *asp-assisted mechanism*. Key interatomic distances of the located transition state are reported in Table S4 of Supporting Information.

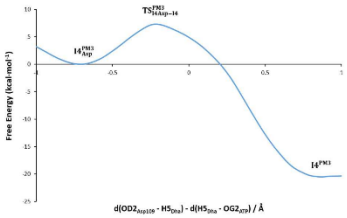


Figure 7. PM3/MM 1D-PMF corresponding to the transition from $I4^{PM3}_{Asp}$ to $I4^{PM3}$.

In order to analyse the impact of the positively charged Arg178B, additional PM3/MM PESs of the most relevant chemical step, from intermediate I2 to intermediate I4, have been computed in both mechanisms by including this residue in the QM sub-set of atoms. Arg178B is interacting with the terminal phosphate oxygen and could preferentially favour the asp-assisted mechanism by stabilizing the O- group in Dha and the γ-phosphate from ATP. A schematic representation of the active site including the residue Arg178B in the QM part, and the obtained PESs are depicted in Figure S11, S12 and S13 of Supporting Information. The resulting PESs show that the inclusion of this residue in the QM region produces a decrease in the energy barrier of the proton transfer from Dha to Asp109A in the *asp-assisted mechanism* but this decrease is not large enough to change the trend obtained with the smaller QM region. Comparing the potential energy barriers of the two mechanism (see Supporting Information) shows that the *substrate-assisted mechanism* (with potential energy barriers of 22 and 10 kcal·mol^{-1}) is still more favourable that for the *asp-assisted mechanism* (with potential energy barriers of 25 and 18 kcal·mol^{-1}).

B3LYP:PM3/MM FESs of the phosphorylation reaction.

According to the PM3/MM FESs, it appears that the *substrate-assisted mechanism* would be kinetically more favourable than the *asp-assisted mechanism*. Then, the correction of the PM3/MM FESs at B3LYP:PM3/MM level has been focused just on the former. The resulting surfaces are presented in Figure 8 while a schematic representation of the mechanism obtained at this level of theory is shown in Scheme 5.

The comparison between these and the previous FESs shown in Figure 3 render interesting conclusions. The first observation is that the B3LYP:PM3/MM results describe the first two steps in the reverse order than the PM3/MM calculations; first, the phosphoryl transfer would take place ($I2^{DFT:PM3}$ to $I3^{DFT:PM3}$) and later the proton from Dha would be transferred to the phosphate group ($I3^{DFT:PM3}$ to $I4^{DFT:PM3}$). The intermediate $I3^{DFT:PM3}$ corresponds to structures where the phosphate is already transferred to the Dha, in contrast to the $I3^{PM3}$. Nevertheless, the distance between the phosphor atom of the transferring phosphate group (PG_{ATP}) and the $O5_{Dha}$ is noticeable larger (1.87 Å) than the final distance measured in, for instance, $I4^{DFT:PM3}$ (1.67 Å).

Another important effect of correcting the PM3/MM FESs is, as observed when comparing Figure 3b with Figure 8b, that the reaction coordinate of the $TS_{I4-P}^{DFT:PM3}$ appears to be displaced to values significantly closer to $P^{DFT:PM3}$ than in the previous TS_{I4-P}^{PM3}. This is valid not only with regard to the proton transfer but also the breaking bond between the Dha and His218A. The transformation from $I4^{DFT:PM3}$ to $P^{DFT:PM3}$ through the intermediate $I5^{DFT:PM3}$ is only slightly different to the description we obtained from the PM3/MM FES (Figure 3b), but with a first TS, $TS_{I4-I5}^{DFT:PM3}$, appearing in a more advance stage of the $NE2_{His218}$-$C1_{Dha}$ breaking bond. Interestingly, different reaction paths computed for the counterpart reaction in solution were also obtained when the QM atoms were described at semiempirical or DFT levels, in our previous QM/MM study of the phosphoryl transfer reaction in solution.[43]

Regarding the free energy barriers, it is unavoidable to stress the dramatic variations of for the phosphoryl transfer step when comparing the PM3/MM with the B3LYP:PM3/MM FESs. While free energy barriers of 19.9 kcal·mol^{-1} and 3.6 kcal·mol^{-1} were obtained for the transformation from $I2^{PM3}$ to $I3^{PM3}$ and from $I3^{PM3}$ to $I4^{PM3}$, respectively, the free energy barriers when the surfaces are corrected at DFT level are 44.3 and 53.6 kcal·mol^{-1}, for the transformation from $I2^{DFT:PM3}$ to $I3^{DFT:PM3}$ and from $I3^{DFT:PM3}$ to $I4^{DFT:PM3}$, respectively. This effect can be derived from the fact that the PM3/MM sampling does not properly cover and explore the representative structures of the B3LYP/MM space. As observed in Scheme 5, the DFT:PM3/MM corrections for the release of the phosphorylated Dha, the transformation from $I4^{DFT:PM3}$ to $P^{DFT:PM3}$, render free energy barriers that are now significantly lower than the original values obtained at PM3/MM level (see Scheme 3). It appears that there are significant biases from the low-level methods indicating a not marginal contribution of degrees of

freedom other than the ones employed in the distinguished reaction coordinate. Then, in order to check the reliability of the correction scheme for this reaction, the next step in our study was to compute PESs at B3LYP/MM level.

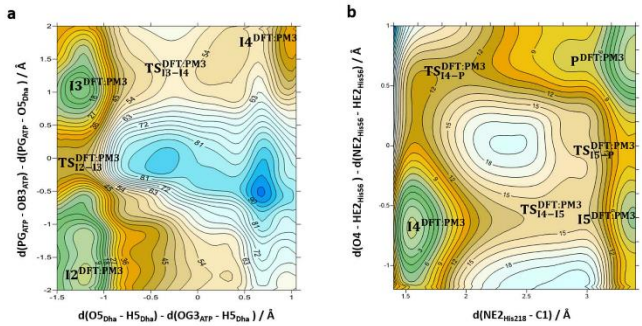
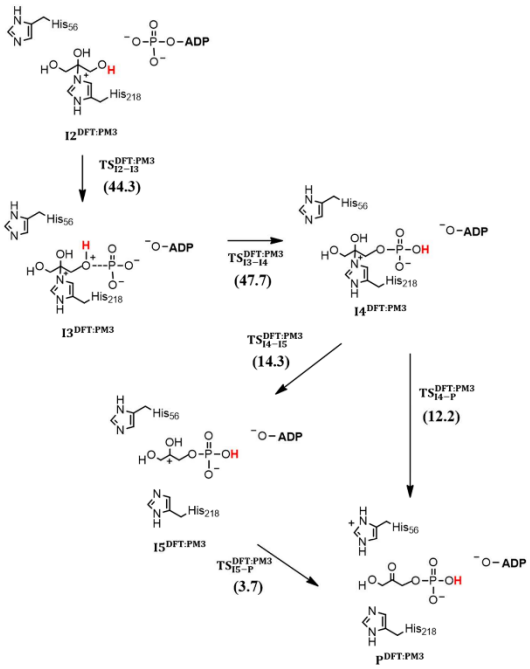


Figure 8. B3LYP:PM3/MM FESs of the *substrate-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E.coli*. Values of the isoenergetic lines are in kcal·mol^{-1}. a) Phosphoryl transfer from ATP to Dha followed by the proton transfer from Dha to the transferred phosphoryl group; and b) releasing of the final products where the proton returns to His56A and the bond with His218A is cleft.

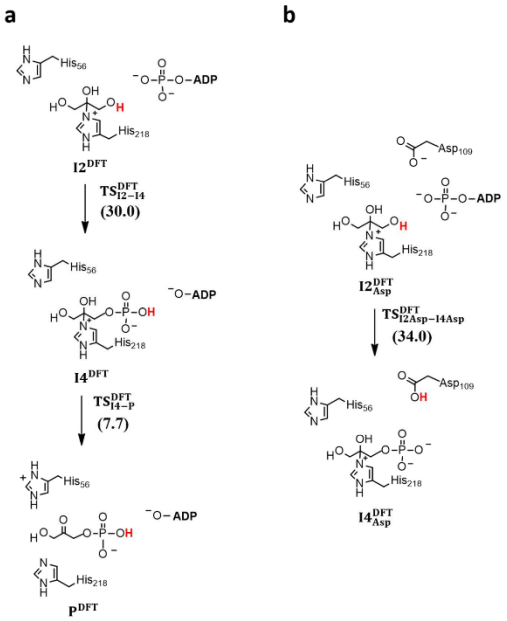


Scheme 5. Schematic representation of the *substrate-assisted mechanism* for the phosphorylation reaction from $I2^{DFT:PM3}$ intermediate to products, $P^{DFT:PM3}$, in DhaK from *E. coli* obtained at DFT:PM3/MM level. Values of free energy barriers, in kcal·mol^{-1}, are reported in brackets.

B3LYP/MM Potential Energy Surfaces.

The B3LYP/MM PESs, deposited as Figures S6 and S7 of the Supporting Information, were used to explore the reaction mechanism and to select initial structures to be refined as TS structures. The nature of the located and characterized TSs were confirmed by computing the Hessian for the QM subset of atom in the presence of the protein and solvent. Then, intrinsic reaction paths (IRCs) were traced down to the reactants and

products valleys to confirm the obtained mechanism. The schematic representation of the mechanisms that are deduced from the B3LYP/MM PESs are shown in Scheme 6, the potential energy profiles are shown in Figure 9, and the TSs of the phosphoryl transfer step of the *substrate-assisted* and *asp-assisted mechanisms* are reported in Figures 10 and 11.



Scheme 6. Schematic representation of the *substrate-assisted* (a) and *asp-assisted* (b) mechanisms for the phosphorylation reaction from ATP to Dha in DhaK from *E.coli* deduced from PESs computed at B3LYP/MM level. Values of potential energy barriers, in kcal·mol^{-1}, are reported in brackets.

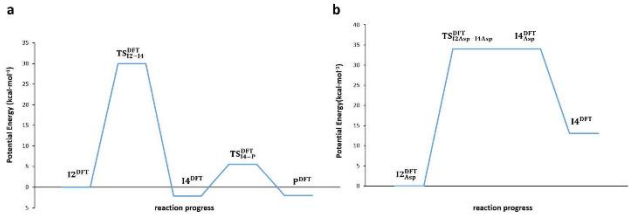


Figure 9. B3LYP /MM potential energy profile of the *substrate-assisted* (a) and the *asp-assisted* (b) phosphoryl transfer reaction mechanisms from ATP to Dha catalyzed by DhaK from *E.coli*.

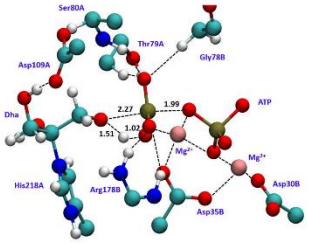


Figure 10. Transition State structure located at B3LYP/MM level, TS_{I2-I4}^{DFT}, corresponding to the phosphoryl transfer step, $I2^{DFT}$ to $I4^{DFT}$, in the *substrate-assisted mechanism*. Dashed lines represent important interactions while key distances are reported in Å.

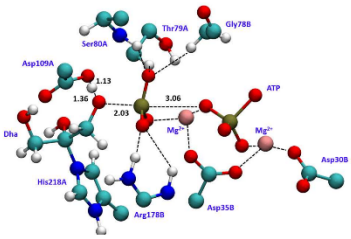


Figure 11. Transition State structure located at B3LYP/MM level, $TS^{DFT}_{I2Asp-I4Asp}$, corresponding to the phosphoryl transfer step, $I2^{DFT}_{Asp}$ to $I4^{DFT}_{Asp}$, in the *asp-assisted mechanism*. Dashed lines represent important interactions while key distances are reported in Å.

The located TSs and their corresponding IRCs confirm that, in both mechanisms, the transformation from $I2^{DFT}$ to $I4^{DFT}$ takes place in a single step, in contrast to the PM3/MM or the DFT:PM3/MM calculations that describe the process through an intermediate I3 (see Scheme 3, 4 and 5). Nevertheless, the comparison of Figure 4 and 6 with Figure 10 and 11 shows that the TSs of the phosphoryl transfer at DFT/MM level are not so different from the average TS structures located at PM3/MM level. Thus, the TS^{DFT}_{I2-I4} located for the *substrate-assisted mechanism* is also describing an asynchronous but concerted process with the phosphoryl transfer in an early stage of the process: the distances from the transferring phosphorous atom to the donor and acceptor atoms are 1.99 and 2.27 Å, respectively. And regarding the *asp-assisted mechanism*, the $TS^{DFT}_{I2Asp-I4Asp}$ is also resembling the TS located at PM3/MM level, $TS^{PM3}_{I3Asp-I4Asp}$. The distances from the transferring phosphorous atom to the donor and acceptor atoms of 3.06 and 2.03 Å, respectively, again describe an asynchronous phosphoryl transfer step where the phosphoryl is in an advance stage of the reaction. A difference is detected with regard to the proton transfer from the Dha to the ATP or the Asp109A, depending on the mechanism. As can be deduced from Figure 10 and 11, the DFT/MM calculations render TSs where the proton is not completely transferred to the acceptor atom, while in the TS structures derived from the PM3/MM calculations the proton was. The angle of $OB3_{ATP}$-PG_{ATP}-$O5_{Dha}$ measured in TS^{DFT}_{I2-I4} and $TS^{DFT}_{I2Asp-I4Asp}$ is 149.1 and 167.2 degrees, respectively, which do not correspond neither to a perfect alignment of

the three atoms but significantly more linear than in the structures deduced from the TSs obtained at PM3/MM level. The distances between the donor and acceptor oxygen atoms, $O5_{Dha}$ – $OB3_{ATP}$, is 4.10 Å for the *substrate-assisted mechanism*, very close to the values deduced at PM3/MM level, but significantly larger, 5.05 Å in the case of *asp-assisted mechanism*. Structural analysis of both TSs are also supporting the role of key residues such as Arg178B, Ser80A and Gly78B, together with the Mg^{2+} cation of the active site. According to the interatomic distances, an interaction is detected between the transferring phosphate and Thr79A but, again, it would correspond to a weak interaction. A table listing the values of the inter-atomic distances in the TSs are reported in Table S5 of the Supporting Information. From the energetic point of view, as observed in Figure 9, the barriers that can be deduced from the B3LYP/MM PESs further support the *substrate-assisted mechanism* by comparison with the *asp-assisted mechanism*: 30 and 34 kcal·mol^{-1}, respectively. In addition, a very exothermic (ca. -25 kcal·mol^{-1}) and barrier-less profile was obtained when trying to connect $I4_{Asp}^{DFT}$ with $I4^{DFT}$ indicating a much more favourable product when the reaction proceeds by means of the *substrate-assisted mechanism*, in agreement with the PM3/MM free energy profile presented in Figure 7. The exploration of the reaction in terms of FESs at DFT level is prohibitive at present, but considering the previous results, smaller values of the free energy barriers would be expected if statistical simulations were carried out.

Finally, according to the DFT/MM results, the conversion from $I4^{DFT}$ into products would be feasible in a thermo-neutral single step (reaction energy equal to 0.2 kcal·mol^{-1}) with a low energy barrier of 7.7 kcal·mol^{-1} (see Scheme 6).

CONCLUSIONS

An exhaustive theoretical study of the phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli* is reported in the present paper. Two mechanisms have been explored according to previous proposals published in the literature for related phosphate transfer reactions. The study has been carried out at three different levels: first calculations of FESs at PM3/MM level; second the semiempirical QM description of the FESs has been improved at DFT level by means of spline corrections; and finally the reaction mechanisms have been explored based on B3LYP/MM PESs. The two explored mechanisms basically differ on the species involved on the activation of the O5 atom of Dha. Thus, in the *substrate-assisted mechanism*, the proton from Dha is

directly transferred to the ATP while in the *asp-assisted mechanism,* Asp109A abstracts the proton from Dha and generates an oxyanion suitable for the attack of the ATP γ-phosphorus atom. Our results suggest that, regardless of the level of theory employed in the calculations, the former would be kinetically more favourable than the later. This conclusion, in agreement with the computational studies of De Vivo et al. for the phosphoryl transfer catalyzed by CDK2,[25] and the early studies on cAMP-PKA by Hart et al.,[27] is also supported by the fact that the intermediate located when exploring the *substrate-assisted mechanism* appears to be significantly more stable than the one obtained when studied the *asp-assisted mechanism*. Our calculations show that after the binding of Dha to the enzyme through a hemiaminal bond with His218A, the phosphate is transferred from ATP to Dha, acting Asp109A not as a base but likely having an important structural role by posing the Dha and other residues of the active site in the proper orientation to stabilize the TS. The Dha would transfer the proton directly to one of the oxygen atoms of the transferring phosphate. Then, a reduction of the activity of the enzyme after mutation of residue Asp109A[34] can not unequivocally demonstrate its role as a base but it can be also indicating an important structural role in the reaction, as we are proposing based on the analysis of the TSs. Thus, considering that Asp109A is not involved in forming or breaking covalent bonds during the *substrate-assisted mechanism*, we suggest that measurements of O^{18} KIEs on the oxygen atoms of Asp109A could report valuable information on the reaction mechanism.

It is important to stress that significant different mechanistic details and energetics are deduced from the different computational methods. Nevertheless, the structures of the TSs corresponding to the phosphate transfer at DFT/MM level significantly resemble the ones obtained at PM3/MM level although subtle nuances can be identified. The TSs suggest that residues such as Gly78B, Thr79A, Ser80A, Arg178B and one of the Mg^{2+} cations of the active site would be stabilizing the transferred phosphate. His218A and His56A would have a decisive role in the preceding chemical steps. This analysis of the role of the key residues of the active site would be partly supported by the kinetic studies of Shi et al.[34], based on activity measurements on site-directed mutants, that indicated the importance of the conserved residues His218A, His56A and Asp109A for the catalytic activity of DhaK.

If kinetics come into focus, dramatic differences are obtained depending on the computational method. Thus, the study based on PM3/MM FESs suggests that the rate

limiting step of the *substrate-assisted mechanism* corresponds to the release of the substrate with a free energy barrier of 21.6 kcal·mol^{-1}, while the phosphor transfer step presents a free energy barrier of 3.6 kcal·mol^{-1}. When the FESs are corrected at B3LYP:PM3/MM level, the free energy barrier for the release of the phosphorylated Dha is decreased to 12.2 kcal·mol^{-1} while the phosphoryl transfer step becomes the rate limiting step with a barrier of 47.7 kcal·mol^{-1}. This dramatic effect can be derived from the limitations of the semiempirical PM3 Hamiltonian that, despite providing reasonable structures for the phosphoryl transfer TSs, the sampling obtained at this level does not properly cover and explore all the representative structures of the B3LYP/MM space. Consequently, the spline corrections technique applied to the PM3/MM FESs can not provide a complete picture of all possible single molecule reaction paths. Nonetheless, the potential energy barriers of the phosphoryl transfer step (coupled to the proton transfer from Dha to the ATP or to the Asp109A) deduced from the B3LYP/MM PESs, which is the rate limiting step, are 30.0 and 34.0 kcal·mol^{-1} for the *substrate-assisted* and *asp-assisted mechanisms*, respectively, thus supporting the former mechanism as the kinetically favoured reaction path. This potential energy barriers appear to be overestimated by comparison with the free energy barrier that can be deduced from experimental rate constants.[34,36] Nevertheless, considering the impact of introducing statistical simulations (when going from the PM3/MM PES to the PM3/MM FESs), an effect on the barrier could also be expected if calculation of PMFs were feasible at DFT/MM level by exploring a huge number of possible different conformations. Finally, since some of the steps of the process involve the transfer of a light particle, the inclusion of quantum tunnelling effects could slightly lower the effective energy barriers. However, a dramatic effect is not expected and, in any case, similar in both mechanism that would not change the observed trend.

SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications website at DOI:

Time evolution of the RMSD of the backbone atoms of the full enzyme and of the residues of the active site along the equilibration MD simulation of the system; time evolution of the distances between ATP and the Mg^{2+} ions, and between His56A and Asp109A with the Dha during the equilibration MD simulations; PESs for all the explored steps at PM3/MM and B3LYP/MM level; PESs from I2 to I4 for both studied

mechanisms including the Arg178B residue in the QM part and potential energy profiles from I2 to I4 for both studied mechanisms including the Arg178B residue in the QM part; key interatomic distances for all the located TSs; and Cartesian coordinates of the QM atoms of the TS corresponding to the phosphoryl transfer located at PM3/MM and B3LYP/MM level.

ACKNOWLEDGEMENTS.

This work was supported by the Spanish Ministerio de Economía y Competitividad (project CTQ2015-66223-C2), Universitat Jaume I (project P1•1B2014-26), Generalitat Valenciana (PROMETEOII/2014/022). V.M. is grateful to the University of Bath for the award of a David Parkin Visiting Professorship. Authors acknowledge computational resources from the Servei d'Informàtica of Universitat Jaume I.

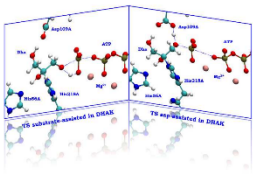
REFERENCES

(1) Hengge, A. C. In *Encyclopedia of Life Sciences*; John Wiley & Sons, London, 2004.
(2) Margutti, S.; Laufer, S. A. *ChemMedChem* **2007**, *2*, 1116-1140.
(3) Smith, G. K.; Ke, Z.; Guo, H.; Hengge, A. C. *J. Phys. Chem. B* **2011**, *115*, 13713-13722.
(4) Kamerlin, S. C. L.; Williams, N. H.; Warshel, A. *J. Org. Chem.* **2008**, *73*, 6960-6969.
(5) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. *Annu. Rev. Biochem.* **2011**, *80*, 669-702.
(6) Ptacek, J.; Snyder, M. *Trends Genet.* **2006**, *22*, 545-554.
(7) Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, *9*, 28-39.
(8) Colinge, J.; César-Razquin, A.; Huber, K.; Breitwieser, F. P.; Májek, P.; Superti-Furga, G. *J. Proteomics* **2014**, *107*, 113-127.
(9) Weinmann, H.; Metternich, R. *ChemBioChem* **2005**, *6*, 455-459.
(10) Bamborough, P. *Expert Opin. Drug Discovery* **2012**, *7*, 1053-1070.
(11) Kamerlin, S. C. L. *J. Org. Chem.* **2011**, *76*, 9228-9238.
(12) Wymore, T.; Field, M. J.; Langan, P.; Smith, J. C.; Parks, J. M. *J. Phys. Chem. B* **2014**, *118*, 4479-4489.
(13) Marcos, E.; Crehuet, R.; Anglada, J. M. *J. Chem. Theory Comput.* **2008**, *4*, 49-63.
(14) Cleland, W. W.; Hengge, A. C. *Chem. Rev.* **2006**, *106*, 3252-3278.
(15) Kamerlin, S. C. L.; Florian, J.; Warshel, A. *ChemPhysChem* **2008**, *9*, 1767-1773.
(16) López-Canut, V.; Roca, M.; Bertran, J.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2010**, *132*, 6955-6963.
(17) Akola, J.; Jones, R. O. *J. Phys. Chem. B* **2003**, *107*, 11774-11783.
(18) Akola, J.; Jones, R. O. *J. Phys. Chem. B* **2006**, *110*, 8121-8129.
(19) Harrison, C. B.; Schulten, K. *J. Chem. Theory Comput.* **2012**, *8*, 2328-2335.
(20) Yang, Y.; Cui, Q. *J. Phys. Chem. A* **2009**, *113*, 12439-12446.
(21) Lu, X. Y.; Ovchinnikov, V.; Demapan, D.; Roston, D.; Cui, Q. *Biochemistry* **2017**, *56*, 1482-1497.
(22) Kiani, F. A.; Fischer, S. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E2947-E2956.
(23) Aqvist, J.; Kamerlin, S. C. L. *Biochemistry* **2015**, *54*, 546-556.
(24) Aqvist, J.; Kamerlin, S. C. L. *ACS Catal.* **2016**, *6*, 1737-1743.
(25) De Vivo, M.; Cavalli, A.; Carloni, P.; Recanatini, M. *Chem. - Eur. J.* **2007**, *13*, 8437-8444.
(26) Ojeda-May, P.; Li, Y.; Ovchinnikov, V.; Nam, K. *J. Am. Chem. Soc.* **2015**, *137*, 12454-12457.
(27) Hart, J. C.; Sheppard, D. W.; Hillier, I. H.; Burton, N. A. *Chem. Commun.* **1999**, 79-80.
(28) Pérez-Gallegos, A.; García-Viloca, M.; González-Lafont, A.; Lluch, J. M. *ACS Catal.* **2015**, *5*, 4897-4912.
(29) Pérez-Gallegos, A.; García-Viloca, M.; González-Lafont, A.; Lluch, J. M. *Phys. Chem. Chem. Phys.* **2015**, *17*, 3497-3511.
(30) Diaz, N.; Field, M. J. *J. Am. Chem. Soc.* **2004**, *126*, 529-542.
(31) Cheng, Y. H.; Zhang, Y. K.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1553-1562.
(32) Jacobsen, D. M.; Bao, Z.-Q.; O'Brien, P.; Brooks, C. L., III; Young, M. A. *J. Am. Chem. Soc.* **2012**, *134*, 15357-15370.
(33) Bastidas, A. C.; Deal, M. S.; Steichen, J. M.; Guo, Y.; Wu, J.; Taylor, S. S. *J. Am. Chem. Soc.* **2013**, *135*, 4788-4798.
(34) Shi, R.; McDonald, L.; Cui, Q.; Matte, A.; Cygler, M.; Ekiel, I. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1302-1307.

(35) Lopata, A.; Jambrina, P. G.; Sharma, P. K.; Brooks, B. R.; Toth, J.; Vertessy, B. G.; Rosta, E. *ACS Catal.* **2015**, *5*, 3225-3237.
(36) García-Alles, L. F.; Siebolo, C.; Nyffeler, T. L.; Flukiger-Bruhwiler, K.; Schneider, P.; Burgi, H. B.; Baumann, U.; Erni, B. *Biochemistry* **2004**, *43*, 13037-13045.
(37) Bachler, C.; Flukiger-Bruhwiler, K.; Schneider, P.; Bahler, P.; Erni, B. *J. Biol. Chem.* **2005**, *280*, 18321-18325.
(38) Zurbriggen, A.; Jeckelmann, J.M.; Christen, S.; Bieniossek, C.; Baumann, U.; Erni, B. *J. Biol. Chem.* **2008**, *283*, 35789-35796.
(39) Enders, D.; Voith, M.; Lenzen, A. *Angew. Chem., Int. Ed.* **2005**, *44*, 1304-1325.
(40) Sánchez-Moreno, I.; Bordes, I.; Castillo, R.; Ruiz-Pernía, J. J.; Moliner, V.; García-Junceda, E. *Int. J. Mol. Sci.* **2015**, *16*, 27835-27849.
(41) Shi, R.; McDonald, L.; Cygler, M.; Ekiel, I. *Struct.* **2014**, *22*, 478-487.
(42) Saier Jr., M. H. *J. Mol. Microbiol. Biotechnol.* **2015**, *25*, 73-78.
(43) Bordes, I.; Ruiz-Pernía, J. J.; Castillo, R.; Moliner, V. *Org. Biomol. Chem.* **2015**, *13*, 10179-10190.
(44) Schaftenaar, G.; Noordik, J. H. *J. Comput. -Aided Mol. Des.* **2000**, *14*, 123-134.
(45) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704-721.
(46) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525-537.
(47) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088-1100.
(48) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666.
(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.
(50) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781-1802.
(51) Vanommeslaeghe, K.; MacKerell, A. D., Jr. *J. Chem. Inf. Model.* **2012**, *52*, 3144-3154.
(52) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. *J. Chem. Inf. Model.* **2012**, *52*, 3155-3168.
(53) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671-690.
(54) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700-733.
(55) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209-220.
(56) Nam, K.; Cui, Q.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 486-504.
(57) Corrie, J. E. T.; Barth, A.; Munasinghe, V. R. N.; Trentham, D. R.; Hutter, M. C. *J. Am. Chem. Soc.* **2003**, *125*, 8546-8554.
(58) Brandt, W.; Dessoy, M. A.; Fulhorst, M.; Gao, W. Y.; Zenk, M. H.; Wessjohann, L. A. *ChemBioChem* **2004**, *5*, 311-323.
(59) Hand, C. E.; Honek, J. F. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 183-188.
(60) Plotnikov, N. V.; Prasad, B. R.; Chakrabarty, S.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **2013**, *117*, 12807-12819.
(61) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199.
(62) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011-1021.
(63) Corchado, J. C.; Coitino, E. L.; Chuang, Y. Y.; Fast, P. L.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 2424-2438.
(64) Nguyen, K. A.; Rossi, I.; Truhlar, D. G. *J. Chem. Phys.* **1995**, *103*, 5522-5530.

(65) Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 1140-1149.
(66) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. *J. Phys. Chem. B* **2004**, *108*, 8427-8433.
(67) Roca, M.; Moliner, V.; Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I. *J. Phys. Chem. A* **2006**, *110*, 503-509.
(68) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S. *J. Phys. Chem. B* **2006**, *110*, 17663-17670.
(69) Renka, R. J. *Siam J. Sci. Stat. Comput.* **1987**, *8*, 393-415.
(70) Renka, R. J. *Acm Transact. Math. Software* **1993**, *19*, 81-94.
(71) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.; Izmaylov, A.; Bloino, J.; Zheng, G.; Sonnenberg, J.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr, J.; Peralta, J.; Ogliaro, F.; Bearpark, M.; Heyd, J.; Brothers, E.; Kudin, K.; Staroverov, V.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.; Iyengar, S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.; Klene, M.; Knox, J.; Cross, J.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; RL, M.; Morokuma, K.; Zakrzewski, V.; Voth, G.; Salvador, P.; Dannenberg, J.; Dapprich, S.; Daniels, A.; Farkas, O.; Foresman, J.; Ortiz , J.; Cioslowski, J. a.; Fox, D. *Gaussian, Inc., Wallingford CT* **2009**.

TOC Graphic



Supporting Information

A Theoretical Study of the Phosphoryl Transfer Reaction from ATP to Dha Catalyzed by DhaK from Escherichia coli

I. Bordes,[1] R. Castillo[1,*] and V. Moliner[1,2,*]

1.Departament de Química Física i Analítica, Universitat Jaume l, 12071Castellón, Spain.

2. Department of Chemistry, University of Bath, Bath BA2 7AY, United Kingdom.

* to whom correspondence should be addressed:

R. Castillo: rcastill@uji.es

V. Moliner: moliner@uji.es

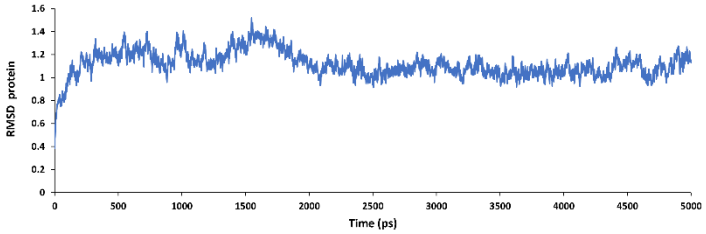


Figure S1: Time evolution of the RMSD of the backbone atoms of the enzyme during 5 ns of MD simulations.

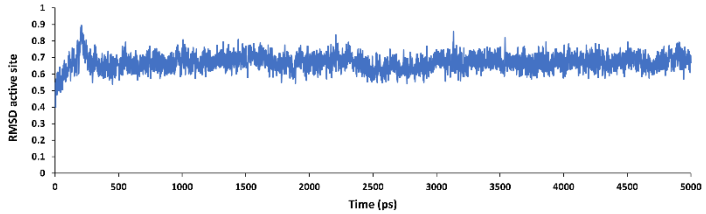


Figure S2: Time evolution of the RMSD computed for the backbone atoms of just the active site residues (those in a radios sphere of 8 Å: Chain A: residues 50-57, 76-81, 105-110, 148-150, 216-219; ChainB: residues 30-40, 76-82, 120-124, 127-133, 176-180, 190-196) during 5 ns of MD simulations.

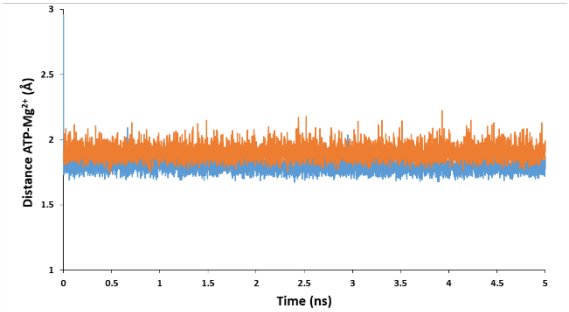


Figure S3. Time evolution of the distance between the Mg^{2+} ion and the oxygen atom OB3 of ATP (blue line) and the other Mg^{2+} ion with the oxygen atom OA1 of ATP (orange line), during the 5 ns of MD simulations.

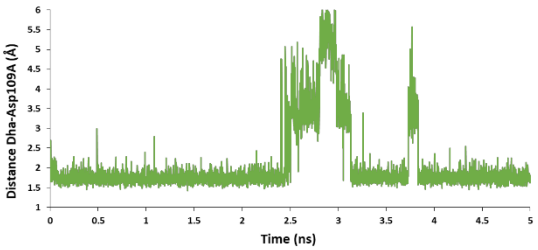


Figure S4. Time evolution of the distance between the hydrogen atom H6 of Dha and the oxygen atom OD1 of Asp109A during the 5 ns of MD simulations.

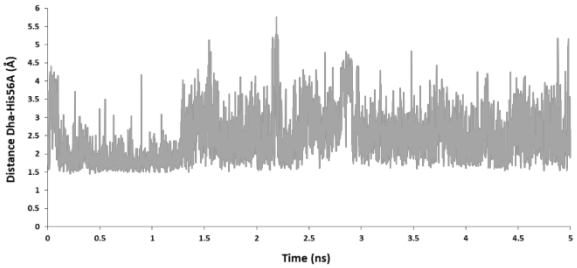


Figure S5. Time evolution of the distance between the hydrogen atom HE2 of His56A and the oxygen atom O4 of Dha during the 5 ns of MD simulations.

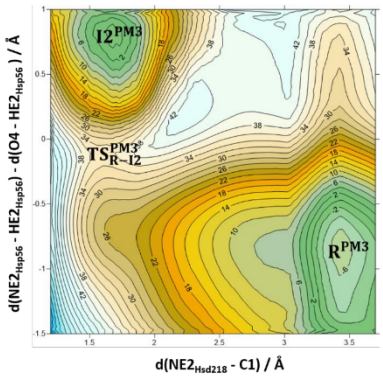


Figure S6: PM3/MM potential energy surface of the first step for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*. Values of the isoenergetic lines on the 2D-PES are in kcal·mol^{-1}.

Table S1. Key interatomic distances with their standard deviations of the transition states located on the PM3/MM PMFs of the first step of the phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*. The values are determined from an average of 100 ps of MD constraining the distinguished reaction coordinate distances.

Distances	TS^{PM3}_{R-I1}(Å)	TS^{PM3}_{I1-I2}(Å)	TS^{PM3}_{R-I2}(Å)
$NE2_{His218A}$ – C1	3.17±0.03	2.32±0.03	1.58±0.03
$NE2_{His56A}$ – $HE2_{His56A}$	1.39±0.04	1.69±0.03	1.22±0.04
O4 – $HE2_{His56A}$	1.14±0.03	1.00±0.02	1.29±0.05
PG_{ATP} – $O5_{Dha}$	5.75±0.30	5.76±0.29	4.31±0.30
PG_{ATP} – $OB3_{ATP}$	1.78±0.04	1.77±0.04	1.80±0.04
$O5_{Dha}$ – $H5_{Dha}$	0.98±0.03	0.98±0.03	0.95±0.02
$OG3_{ATP}$ – $H5_{Dha}$	5.67±0.55	5.86±0.27	3.16±0.70
$OD2_{Asp109A}$ – $H5_{Dha}$	1.91±0.55	1.77±0.11	3.93±0.98
$OB3_{ATP}$ – $H5_{Dha}$	7.29±0.46	7.37±0.25	5.30±0.67
$OG2_{ATP}$ – $H5_{Dha}$	5.03±0.39	5.01±0.25	3.93±0.58

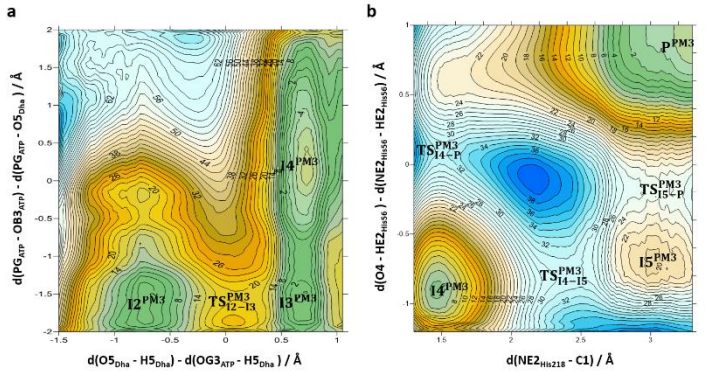


Figure S7. PM3/MM potential energy surfaces of the chemical steps of the *substrate-assisted mechanism* for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*. a) Proton transfer from Dha to the transferring phosphoryl group of ATP followed by the phosphoryl transfer to Dha; and b) releasing of the final products where the proton returns to His56A and the bond with His218A is cleft. Values of the isoenergetic lines on the 2D-PES are in kcal·mol^{-1}.

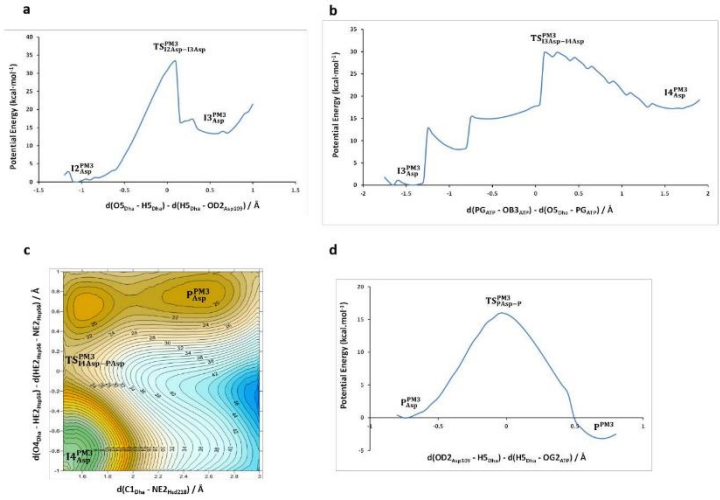


Figure S8. PM3/MM potential energy surfaces of the chemical steps of the *asp-assisted mechanism* for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*. a) Proton transfer from Dha to Asp109A; b) phosphoryl transfer from ATP to Dha; c) releasing of the products where the proton returns to His56A and the bond with His218A is cleft; and d) proton transfer from Asp109A to ADP, reaching the final products. Values of the isoenergetic lines on the 2D-PES are in kcal·mol^{-1}.

Table S2. Key interatomic distances with their standard deviations of the transition states located on the PM3/MM PMFs of the chemical steps of the *substrate-assisted mechanism* for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*. The values are determined from an average of 100 ps of MD constraining the distinguished reaction coordinate distances.

Distances	TS^{PM3}_{I2-I3}(Å)	TS^{PM3}_{I3-I4}(Å)	TS^{PM3}_{I4-I5}(Å)	TS^{PM3}_{I5-P}(Å)	TS^{PM3}_{I4-P}(Å)
$NE2_{His218A}$ – C1	1.53±0.04	1.53±0.04	2.40±0.03	3.01±0.03	1.58±0.03
$NE2_{His56A}$ – $HE2_{His56A}$	1.92±0.27	5.04±0.45	1.72±0.03	1.45±0.04	1.21±0.04
O4 – $HE2_{His56A}$	0.97±0.03	0.97±0.03	1.00±0.02	1.13±0.03	1.36±0.05
PG_{ATP} – $O5_{Dha}$	3.48±0.04	2.57±0.05	1.78±0.04	1.78±0.04	1.75±0.04
PG_{ATP} – $OB3_{ATP}$	1.77±0.03	1.77±0.04	3.64±0.14	3.72±0.14	3.83±0.19
$O5_{Dha}$ – $OB3_{ATP}$	4.58±0.09	3.98±0.09	5.09±0.13	5.23±0.16	5.05±0.14
$O5_{Dha}$ – $H5_{Dha}$	1.14±0.03	1.72±0.03	2.64±0.14	2.78±0.16	3.23±0.30
$OG3_{ATP}$ – $H5_{Dha}$	1.17±0.04	0.97±0.02	0.95±0.02	0.95±0.02	0.96±0.02
$OD2_{Asp109A}$ – $H5_{Dha}$	5.54±0.27	6.07±0.27	6.18±0.23	6.31±0.19	6.46±0.25
$OB3_{ATP}$ – $H5_{Dha}$	3.46±0.08	3.07±0.07	3.17±0.14	3.18±0.17	2.27± 0.64
$OG2_{ATP}$ – $H5_{Dha}$	3.14±0.12	2.89±0.14	2.94±0.21	2.82±0.22	2.79±0.27
$OG2_{ATP}$ – $HG1_{Thr79A}$	3.19±0.38	2.20±0.34	4.47±0.25	4.55 ±0.27	5.23±0.33
$OG2_{ATP}$ – HA_{Thr79A}	2.69±0.16	2.61±0.15	2.62±0.15	2.69±0.16	2.91±0.24
$OG2_{ATP}$ – H_{Ser80A}	2.03±0.15	1.97±0.14	1.78±0.10	1.81±0.11	1.93±0.34
$OG2_{ATP}$ – $HA2_{Gly78B}$	2.61±0.19	2.56±0.17	2.54±0.14	2.61±0.17	3.05±0.38
$OG3_{ATP}$ – $HE_{Arg178B}$	3.65±0.26	3.24±0.57	3.87±0.22	4.00±0.21	4.17±0.28
$OG3_{ATP}$ – MG	1.86±0.03	1.86±0.03	2.48±0.04	2.49±0.04	2.49±0.04

Table S3. Key interatomic distances with their standard deviations of the transition states located on the PM3/MM PMFs of the chemical steps of the *asp-assisted mechanism* for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*. The values are determined from an average of 100 ps of MD constraining the distinguished reaction coordinate distances.

Distances	$TS^{PM3}_{I2Asp-I3Asp}$(Å)	$TS^{PM3}_{I3Asp-I4Asp}$(Å)	$TS^{PM3}_{I4Asp-PAsp}$(Å)	TS^{PM3}_{PAsp-P}(Å)
NE2$_{His218}$ – **C1**	1.52±0.04	1.52±0.03	1.57±0.03	3.02±0.17
NE2$_{His56A}$ – **HE2**$_{His56A}$	3.77±0.57	1.95±0.30	1.22±0.04	1.01±0.03
O4 – **HE2**$_{His56A}$	0.95±0.02	0.97±0.03	1.34±0.05	1.91±0.25
PG$_{ATP}$ – **O5**$_{Dha}$	5.15±0.25	1.98±0.06	1.79±0.04	1.77±0.04
PG$_{ATP}$ – **OB3**$_{ATP}$	1.78±0.04	2.32±0.06	3.55±0.12	3.55±0.12
O5$_{Dha}$ – **OB3**$_{ATP}$	6.82±0.24	4.07±0.11	5.14±0.12	5.12±0.12
O5$_{Dha}$ – **H5**$_{Dha}$	1.18±0.04	1.74±0.06	3.03±0.23	2.83±0.13
OG3$_{ATP}$ – **H5**$_{Dha}$	5.13±0.24	4.03±0.12	4.17±0.12	3.71±0.09
OD2$_{Asp109A}$ – **H5**$_{Dha}$	1.14±0.04	0.99±0.03	0.98±0.03	1.11±0.03
OB3$_{ATP}$ – **H5**$_{Dha}$	6.77±0.23	5.10±0.18	5.82±0.15	5.20±0.15
OG2$_{ATP}$ – **H5**$_{Dha}$	4.80±0.26	3.02±0.23	1.74±0.07	1.20±0.04
OG2$_{ATP}$ – **HA**$_{Thr79A}$	2.64±0.14	2.90±0.22	3.45±0.19	3.43±0.18
OG2$_{ATP}$ – **H**$_{Ser80A}$	1.90±0.11	1.89±0.16	2.65±0.42	3.51±0.39
OG2$_{ATP}$ – **HA2**$_{Gly78B}$	2.56±0.17	2.69±0.21	4.03±0.27	4.14±0.29
OG3$_{ATP}$ – **HA**$_{Thr79A}$	2.47±0.12	2.53±0.13	2.54±0.13	2.63±0.15
OG3$_{ATP}$ – **MG**	1.85±0.03	1.86±0.03	1.87±0.03	1.87±0.03
OG3$_{ATP}$ – **HH11**$_{Arg178B}$	1.95±0.11	2.39±0.30	3.43±0.14	3.53±0.12
OG1$_{ATP}$ – **HH12**$_{Arg178B}$	1.86±0.13	2.09±0.23	3.43±0.02	3.56±0.26

Table S4. Key interatomic distances with their standard deviations of the transition state located on the PM3/MM PMF of the chemical step corresponding to the transition from $I4^{PM3}_{Asp}$ to $I4^{PM3}$ for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*. The values are determined from an average of 100 ps of MD constraining the distinguished reaction coordinate distances.

Distances	$TS^{PM3}_{I4Asp-I4}$(Å)
$NE2_{His218A}$ – C1	1.52±0.03
$NE2_{His56A}$ – $HE2_{His56A}$	3.78±0.70
O4 – $HE2_{His56A}$	0.96±0.02
PG_{ATP} – $O5_{Dha}$	1.93±0.07
PG_{ATP} – $OB3_{ATP}$	1.92±0.07
$O5_{Dha}$ – $H5_{Dha}$	2.82±0.16
$OG3_{ATP}$ – $H5_{Dha}$	3.38±0.13
$OD2_{Asp109A}$ – $H5_{Dha}$	1.34±0.05
$OB3_{ATP}$ – $H5_{Dha}$	3.82±0.11
$OG2_{ATP}$ – $H5_{Dha}$	1.07±0.03

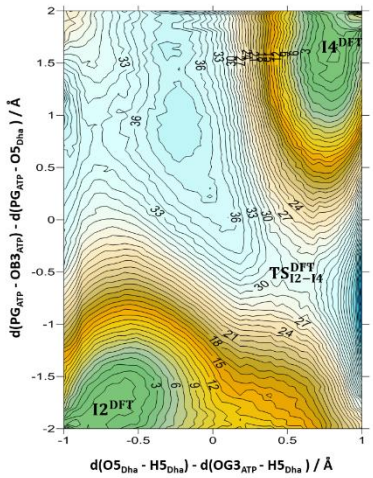


Figure S9. B3LYP/MM potential energy surface of the *substrate-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*. Values of the isoenergetic lines are in kcal·mol^{-1}.

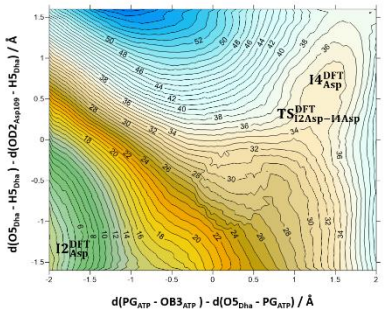


Figure S10. B3LYP/MM potential energy surface of the *asp-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*. Values of the isoenergetic lines are in kcal·mol^{-1}.

Table S5. Key interatomic distances of the transition states located at B3LYP/MM level of the chemical steps of the *substrate-assisted* and *asp-assisted* phosphoryl transfer reactions from ATP to Dha catalyzed by DhaK from *E. coli*.

Distances	TS^{DFT}_{I2-I4}(Å)	$TS^{DFT}_{I2Asp-I4Asp}$(Å)
$NE2_{His218A}$ – C1	1.55	1.55
$NE2_{His56A}$ – $HE2_{His56A}$	1.88	1.97
O4 – $HE2_{His56A}$	1.00	1.00
$O5_{Dha}$ – $H5_{Dha}$	1.51	1.36
$OD2_{Asp109A}$ – $H5_{Dha}$	6.82	1.13
$OG3_{ATP}$ – $H5_{Dha}$	1.02	3.91
PG_{ATP} – $OB3_{ATP}$	1.99	3.06
PG_{ATP} – $O5_{Dha}$	2.27	2.03
$O5_{Dha}$ – $OB3_{ATP}$	4.10	5.05
$OB3_{ATP}$ – $H5_{Dha}$	3.35	5.74
$OG2_{ATP}$ – $H5_{Dha}$	2.99	2.98
$OG2_{ATP}$ – HA_{Thr79A}	2.72	3.00
$OG2_{ATP}$ – H_{Ser80A}	1.85	1.85
$OG2_{ATP}$ – $HA2_{Gly78B}$	2.36	2.45
$OG1_{ATP}$ – $HH12_{Arg178B}$	1.70	1.56
$OG1_{ATP}$ – $HE_{Arg178B}$	2.34	2.72
$OG3_{ATP}$ – MG	2.00	2.10

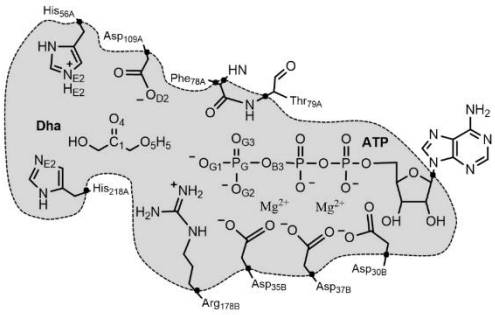


Figure S11. Schematic representation of the active site of DhaK adding the residue Arg178B. Grey region contains atoms treated quantum mechanically. Link atoms are represented as black dots.

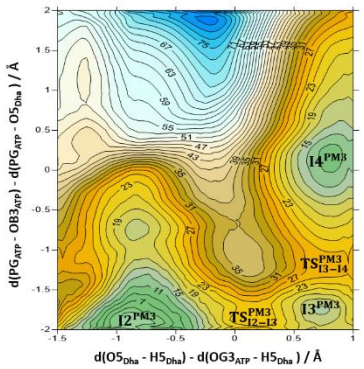


Figure S12. PM3/MM potential energy surface for the transformations from $I2^{PM3}$ to $I4^{PM3}$ following the *substrate-assisted mechanism* catalyzed by DhaK from *E. coli*, including Arg178B in the QM part. Values of the isoenergetic lines on the 2D-PES are in kcal·mol^{-1}.

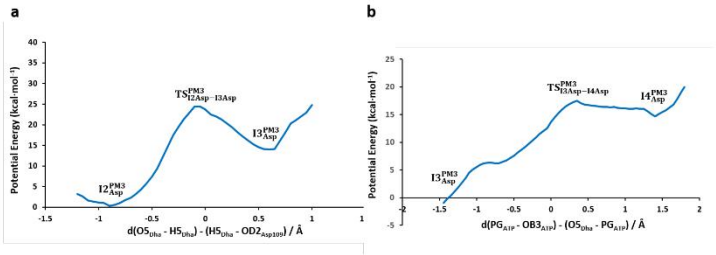


Figure S13. PM3/MM potential energy surfaces of the more relevant chemical steps of the *asp-assisted mechanism* for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*, including Arg178B in the QM part. a) Proton transfer from Dha to Asp109A; and b) phosphoryl transfer from ATP to Dha.

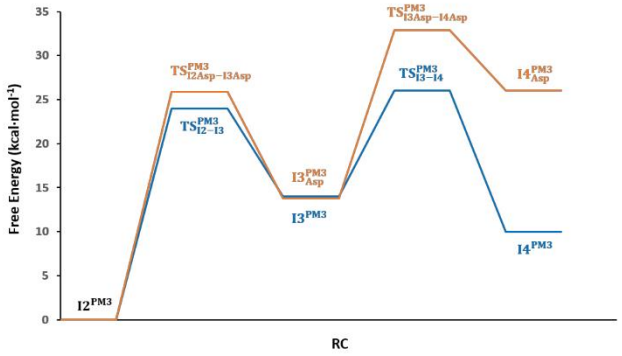


Figure S14. PM3/MM potential energy profiles of the more relevant chemical steps of the *substrate-assisted* (blue line) and *asp-assisted* (orange line) mechanisms for the phosphorylation reaction from ATP to Dha catalyzed by DhaK from *E. coli*, including Arg178B in the QM part.

Table S6. Cartesian coordinates of the QM atoms of the $TS^{PM3}_{I2Asp-I4Asp}$ localized at PM3/MM level of the *asp-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*.

Dha			
C1	-2.9988603821	-1.2234492628	-3.6708257192
C2	-1.6396761118	-0.7676623245	-2.9132391705
H21	-1.6314924817	0.3582326051	-2.9447642152
H22	-1.8016050591	-1.0334340806	-1.8410418178
C3	-4.2610222821	-0.7527367361	-2.8602341638
H31	-5.1992260049	-0.9181533019	-3.4262643579
H32	-4.3315556082	-1.3240846861	-1.9083713559
O4	-3.0150208217	-0.7173300779	-4.9628625903
O5	-0.5431392252	-1.2792507240	-3.4503144805
H5	-1.1354919431	1.5452083559	-1.7860681227
O6	-4.2750771085	0.6235034134	-2.6292419065
H6	-3.6548773167	0.8188780072	-1.9240687791
ATP			
C1'	9.4473643384	-5.5320981206	2.9358564068
H1'	10.4009214788	-4.9691286648	2.9676380863
C2'	9.6470716316	-7.0610903343	2.8456047613
H2'	11.1375381222	-8.0060574060	2.1561348735
O2'	11.0244785236	-7.3598270796	2.8579308166
HT2	9.1940172998	-7.5991295805	3.7077190516
C3'	8.9867208801	-7.4649172795	1.4904539664
H3'	7.9991372555	-7.9584614565	1.6351325016
O3'	9.8491150296	-8.3144487408	0.7687344174
HT3	9.9087885196	-9.1417412273	1.2526971091
C4'	8.8508431357	-6.1146285588	0.7367783984
O4'	8.7679623410	-5.0965699152	1.7478345433
H4'	9.7732462121	-5.9021711442	0.1434316939
C5'	7.6062312535	-6.0428899342	-0.1881871455
H51	6.7177233700	-5.6759707554	0.3615478701
H52	7.3678850477	-7.0609341177	-0.5573208654
O5'	7.9065119561	-5.2069119852	-1.2566274753
PA	6.9051530086	-4.7829796009	-2.6198493337
OA1	6.4238540779	-6.0475892106	-3.4413148337
OA2	7.9692162152	-3.8900025855	-3.2817516752
OA3	5.6222835346	-3.9240513747	-1.9906524996
PB	3.8722802508	-4.3339359432	-2.1362190544
OB1	3.8383441951	-5.5186139035	-3.2366919363
OB2	3.4736947006	-4.6750215743	-0.7062267417
OB3	3.1099856333	-3.0596878090	-2.7650153611
PG	1.0866799294	-1.9392292542	-2.6496218904
OG1	0.3730636326	-2.7676756846	-1.5488212215
OG2	1.6072346901	-0.5448317439	-2.2525520916
OG3	1.4051500422	-2.5421272940	-4.0878185578

His_{56A}			
CB	-8.3000067826	-0.3002185048	-7.9937392740
HB1	-9.1226795324	-0.2008493142	-7.2611478067
HB2	-7.9324108774	0.7221274916	-8.1900603695
CG	-7.2160343925	-1.0562327561	-7.3339613174
ND1	-6.9846299525	-2.4382533334	-7.3994838501
HD1	-7.5543755642	-3.0895614407	-7.8820209382
CE1	-5.8638411175	-2.7368837342	-6.6499047527
HE1	-5.3988395953	-3.7176479825	-6.5234522869
NE2	-5.3935764339	-1.5800705335	-6.1038084258
HE2	-3.8646351075	-0.9563716086	-5.3800360032
CD2	-6.2225731156	-0.5412057401	-6.5087933113
HD2	-6.0486185939	0.4864482982	-6.1952945849
His_{218A}			
CB	-2.5218112321	-6.1325793426	-5.1820682716
HB1	-1.6489636931	-5.9701165438	-5.8434091833
HB2	-2.1698662810	-6.8820236463	-4.4496103491
CG	-2.8002737264	-4.8655012336	-4.4697585290
ND1	-3.4852003482	-4.8440938833	-3.2467243142
HD1	-3.7742902904	-5.6552178466	-2.7482693266
CE1	-3.6470737530	-3.5428972971	-2.8363241682
HE1	-4.1667286812	-3.1638451272	-1.9408690712
NE2	-3.0354217596	-2.7445188643	-3.7796407504
CD2	-2.5053996761	-3.5422902277	-4.7989685123
HD2	-1.9299475381	-3.1316064758	-5.6550805307
Asp_{109A}			
CB	-1.4821651273	2.5268802260	1.2758638602
HB1	-2.1709100080	3.3685298977	1.4942388101
HB2	-1.7875374405	1.7520795547	2.0080876826
CG	-1.8782374936	2.0317367739	-0.0801294948
OD1	-2.9631442332	1.5763530322	-0.4072059869
OD2	-0.9482203775	2.1447250892	-1.0529084296
MG			
MG	4.8851230679	-6.5306254669	-4.4296073954
MG	2.8166787599	-3.6753610180	-4.4850629358
Asp_{30B}			
CB	6.6832707798	-9.9504853892	-5.4748606317
HB1	6.7927061005	-9.8740515172	-6.5681342773
HB2	7.6761084796	-9.6222385431	-5.1072143080
CG	5.7444731521	-8.8382867279	-4.9766654362
OD1	4.9135468492	-8.9801355635	-4.0624351694
OD2	5.7832563968	-7.7117236258	-5.5840970627
Asp_{35B}			
CB	0.9719351055	-7.2195801025	-5.8110833004
HB1	0.1443479990	-7.0416777149	-5.0936001854
HB2	0.5857476136	-6.8011090795	-6.7580510927
CG	2.1538356996	-6.3271190215	-5.3533066198
OD1	1.8372129098	-5.1353126229	-5.1113700412
OD2	3.2814778692	-6.9328536248	-5.3360970560

Asp_{37B}			
CB	4.6347354934	-4.3156733954	-7.8299703357
HB1	3.6138047436	-4.3746133169	-8.2586458236
HB2	5.0192905760	-3.3529989143	-8.2199886682
CG	4.3978099267	-4.1159912548	-6.3461245594
OD1	4.7412702153	-4.9411545144	-5.4597473570
OD2	3.8644143614	-3.0323379397	-5.9010627329
Phe_{78A}			
CA	-0.2757833650	-2.4856911772	-7.7296901130
HA	-1.1051533910	-2.1496281373	-7.0569175308
C	1.0034260953	-2.0412770163	-7.0575238704
O	1.9315355480	-2.7228685555	-6.6418432871
Thr_{79A}			
N	1.2156084164	-0.6241081486	-6.9412039234
H	0.3559358833	-0.1279634226	-7.0708652467

Table S7. Cartesian coordinates of the QM atoms of the TS_{I2-I4}^{PM3} localized at PM3/MM level of the *substrate-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*.

Dha			
C1	-2.8265367275	-1.0594915193	-3.8733540018
C2	-1.3190998039	-0.5770846503	-3.5266256005
H21	-1.3078582628	0.5093569082	-3.7580783646
H22	-1.2394131379	-0.6853762038	-2.4198305367
C3	-3.7997916126	-0.4135275918	-2.8225853804
H31	-4.8602370111	-0.6213491479	-3.0710337136
H32	-3.5885801790	-0.8281633350	-1.8177339707
O4	-3.1276138707	-0.7155635709	-5.1801759132
O5	-0.3569620667	-1.2110610930	-4.1635467064
H5	0.6103180663	-2.6158076707	-4.5068948596
O6	-3.6891259838	0.9771198538	-2.8176893275
H6	-3.1999520949	1.2467366841	-2.0185072758
ATP			
C1'	9.4092189190	-5.4876166459	2.8950834103
H1'	10.3383489150	-4.8825777579	2.9239161667
C2'	9.6890586719	-7.0093982247	2.7888822496
H2'	11.2754419620	-8.0153571748	2.3888210543
O2'	11.0875381231	-7.1839176802	2.8072464285
HT2	9.2717632945	-7.5754388632	3.6529016660
C3'	8.9946414257	-7.4312442950	1.4563083582
H3'	7.9973652191	-7.8831436852	1.6712583178
O3'	9.7314201616	-8.3198243556	0.6572133151
HT3	9.8432826798	-9.1269780380	1.1639437195

C4'	8.8424416515	-6.0859281095	0.6975290789
O4'	8.7174630058	-5.0717750331	1.7060981144
H4'	9.7673789075	-5.8569738682	0.1141769856
C5'	7.6069613295	-6.0553187017	-0.2339531125
H51	6.7104429647	-5.6773906631	0.2944804016
H52	7.3812169603	-7.0819391281	-0.5852196541
O5'	7.9057009667	-5.2431439020	-1.3269681090
PA	6.8915151168	-4.8496759323	-2.6590482436
OA1	6.3943466431	-6.1144803430	-3.4735061340
OA2	7.8612506026	-3.8843638075	-3.3241208809
OA3	5.5381070079	-4.1019898179	-1.9919146221
PB	3.8729128902	-4.6212767687	-2.0810542778
OB1	3.7724710603	-5.7700807738	-3.1733647772
OB2	3.4243272157	-4.8245726961	-0.6671520794
OB3	2.9982067831	-3.3535905594	-2.7987042439
PG	1.6554268241	-2.2131805689	-2.4780521366
OG1	0.6967391874	-2.8718338692	-1.5177118655
OG2	2.1393186785	-0.8216725069	-2.2884309457
OG3	1.4672621441	-2.7360401230	-4.0916849796
His_{56A}			
CB	-8.3143483860	-0.2530601903	-8.0381239333
HB1	-9.1423316496	-0.0615321864	-7.3310437984
HB2	-7.8802849498	0.7341167254	-8.2777623150
CG	-7.2995442605	-1.0366404795	-7.3075860805
ND1	-7.0784878912	-2.4193671806	-7.3896161322
HD1	-7.6146355717	-3.0502496195	-7.9337864167
CE1	-6.0377863001	-2.7545326541	-6.5455237994
HE1	-5.6026527245	-3.7450336718	-6.3994535091
NE2	-5.6108897008	-1.6185601811	-5.9270943461
HE2	-4.0727138154	-0.8709047360	-5.3414021558
CD2	-6.3804227711	-0.5553420357	-6.3814659685
HD2	-6.2126204454	0.4639424745	-6.0343215838
His_{218A}			
CB	-2.5046133787	-6.0530150022	-5.1354166790
HB1	-1.6194085755	-5.9244808452	-5.7857162750
HB2	-2.1789045024	-6.7769026086	-4.3649833968
CG	-2.7733492394	-4.7499564077	-4.4756348913
ND1	-3.4946008127	-4.6524004612	-3.2773179291
HD1	-3.8322995204	-5.4279820733	-2.7584966505
CE1	-3.6023692695	-3.3279149829	-2.9059549087
HE1	-4.1500812437	-2.9276868223	-2.0401382574
NE2	-2.9361679267	-2.5870061071	-3.8529865338
CD2	-2.4194091573	-3.4488018362	-4.8305919331
HD2	-1.8252001311	-3.0757646936	-5.6900534465
Asp_{109A}			
CB	-1.5470576654	2.7366689630	1.4319640017
HB1	-2.2319861195	3.5583299297	1.7054265824
HB2	-1.8782352416	1.8892745353	2.0526072784
CG	-1.8688625657	2.4372245332	-0.0442684979

OD1	-2.8563733326	1.7194415611	-0.3483817934
OD2	-1.1255847007	2.9387042799	-0.9286309768
MG			
MG	4.8650071327	-6.6484047297	-4.4268401680
MG	2.7230097035	-4.0358223829	-4.5280830157
Asp_{30B}			
CB	6.6830388752	-10.0104585659	-5.4829121910
HB1	6.8145054810	-9.9327526379	-6.5745034098
HB2	7.6635868460	-9.6727436148	-5.0931131380
CG	5.7089388559	-8.9411071323	-5.0018961241
OD1	4.8851461192	-9.0829760799	-4.0796599462
OD2	5.6631602840	-7.8264075552	-5.6412960795
Asp_{35B}			
CB	0.8787749399	-7.3447644392	-5.8086624724
HB1	-0.0125252584	-7.2883484830	-5.1416298376
HB2	0.4875106438	-6.9618226106	-6.7722827925
CG	1.8417301472	-6.2896770881	-5.2856401103
OD1	1.3272264710	-5.1519212828	-5.0208095373
OD2	3.0813162280	-6.4422134809	-5.0956791621
Asp_{37B}			
CB	4.6596547246	-4.3301621395	-7.8632145972
HB1	3.6654664950	-4.3446784040	-8.3545823218
HB2	5.0765312235	-3.3526673046	-8.1713042683
CG	4.3538508376	-4.2564607132	-6.3887628739
OD1	4.6599401393	-5.1805906660	-5.5862536526
OD2	3.7608898279	-3.2473992127	-5.8651405493
Phe_{78A}			
CA	-0.2704337256	-2.5501934155	-7.7861180195
HA	-1.0925695678	-2.2122528395	-7.1047425192
C	0.9991097601	-2.1142417812	-7.0931978491
O	1.8615582118	-2.8278420678	-6.5769932008
Thr_{79A}			
N	1.2686330934	-0.7104658034	-7.0558049031
H	0.4485312651	-0.1974721389	-7.3189758348

Table S8. Cartesian coordinates of the QM atoms of the $TS^{DFT}_{I2Asp-I4Asp}$ localized at B3LYP/MM level of the *asp-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*.

Dha			
C1	-3.0333609231	-1.3149799014	-3.4958818440
C2	-1.8706761787	-1.2216844609	-2.4274281111
H21	-2.3555549952	-0.8670638293	-1.5044230070
H22	-1.5544403654	-2.2457484447	-2.2493075209
C3	-4.3240137243	-0.6413450656	-2.9672689777
H31	-5.1345123953	-0.8242603305	-3.6799295522
H32	-4.6119363091	-1.0903337132	-2.0019344506
O4	-2.7332067924	-0.8483104132	-4.7624035698

O5	-0.7306518767	-0.4549530761	-2.6932681950
H5	-0.6956718902	0.5036524574	-1.7282688974
O6	-4.1341831232	0.7477626965	-2.8725439155
H6	-3.5942982370	0.9557080869	-2.0768212278
ATP			
C1'	9.2611321409	-5.5215486328	3.0096757890
H1'	10.1889531535	-4.9347294732	3.0415941248
C2'	9.6136077524	-7.0138394659	2.9059649792
H2'	11.1638331070	-7.9008286906	2.2043279319
O2'	11.0203616071	-7.1470060925	2.8075056798
HT2	9.2355743441	-7.5877878530	3.7670077656
C3'	8.9128372238	-7.4608125053	1.5801121887
H3'	7.9258885322	-7.8838424739	1.8168156380
O3'	9.7034823360	-8.3427239285	0.8262472466
HT3	9.7015843225	-9.2730948463	1.2092241568
C4'	8.7459816612	-6.1284370199	0.8265616098
O4'	8.4666103762	-5.1756675180	1.8686628822
H4'	9.7078332165	-5.8914516416	0.3466227444
C5'	7.6503257451	-6.0789850961	-0.2278873028
H51	6.6685415814	-6.0131363020	0.2626367975
H52	7.6812731703	-6.9962656096	-0.8261420498
O5'	7.8824583566	-4.9411550338	-1.0516745115
PA	7.0311587281	-4.6714716399	-2.4508950237
OA1	6.6851301674	-6.0050439962	-3.0946044953
OA2	7.9400742392	-3.7247929771	-3.2046835243
OA3	5.7108595867	-3.9142497042	-1.9263192304
PB	4.0404855455	-4.4361170931	-1.8930364254
OB1	3.9816603845	-5.8140477672	-2.5882664308
OB2	3.6539911169	-4.4004639415	-0.4411782030
OB3	3.4067927679	-3.3529181452	-2.8137179387
PG	0.9057837643	-1.6309492338	-2.4701754448
OG1	0.5059738235	-2.2251605444	-1.1352529421
OG2	1.9637541608	-0.5781136887	-2.5312927455
OG3	0.7453271206	-2.5419865662	-3.7080653496
His_{56A}			
CB	-8.2807206595	-0.2599580440	-8.0101815836
HB1	-9.0940702389	-0.2229481044	-7.2736085626
HB2	-7.9911317916	0.7761758757	-8.2017125577
CG	-7.1068312601	-0.9658778405	-7.3936867531
ND1	-6.8813146808	-2.3272221862	-7.4341630648
HD1	-7.4930981612	-3.0058311191	-7.8753152385
CE1	-5.7440508516	-2.5944493859	-6.7402691624
HE1	-5.3468738480	-3.5936122466	-6.6431260062
NE2	-5.2131771513	-1.4841912397	-6.2420735524
HE2	-3.5634287766	-0.9674528811	-5.3022957614
CD2	-6.0718502778	-0.4732203244	-6.6348734103
HD2	-5.9082154086	0.5490051549	-6.3355374403
His_{218A}			
CB	-2.5774774719	-6.0771063094	-5.2765785164

HB1	-1.7477374175	-5.8402177124	-5.9419612904
HB2	-2.2082594689	-6.8560278284	-4.6070640363
CG	-2.9181123304	-4.8694258498	-4.4716096600
ND1	-3.5417003776	-4.9562761591	-3.2409529490
HD1	-3.7894583980	-5.8858705852	-2.8292710338
CE1	-3.7463146715	-3.7176448196	-2.7570143422
HE1	-4.2327071572	-3.4759974887	-1.8264031998
NE2	-3.2706248014	-2.8363423011	-3.6446883440
CD2	-2.7380928919	-3.5382003452	-4.7049363972
HD2	-2.2424343358	-3.0286191263	-5.5061220809
Asp_{109A}			
CB	-1.3491919606	2.3607973184	1.1438664892
HB1	-2.0585360119	3.1950855710	1.1839940641
HB2	-1.6847017839	1.6728570149	1.9289414452
CG	-1.6096027877	1.6462386066	-0.1994727404
OD1	-2.7918988043	1.4872546408	-0.5415147095
OD2	-0.5729722925	1.2110249803	-0.8529381048
MG			
MG	4.9904978399	-6.5674737574	-4.1523470269
MG	2.3987358593	-3.5861867543	-4.4679034147
Asp_{30B}			
CB	6.9541412013	-10.0486565329	-5.5249180443
HB1	6.9560691568	-9.9678913160	-6.6158796947
HB2	7.9799266251	-9.8041818968	-5.2196603346
CG	6.0527163539	-8.9425564110	-4.9080191135
OD1	5.5321534905	-9.0883078424	-3.7869664757
OD2	5.9323617580	-7.8571433613	-5.5885505717
Asp_{35B}			
CB	0.7606740638	-7.5123422492	-5.6801746174
HB1	-0.0042992126	-7.4952319176	-4.8957870163
HB2	0.3012686706	-7.0094438698	-6.5379438756
CG	1.9297338598	-6.6205256158	-5.1747215498
OD1	1.6260265558	-5.4005027707	-4.9944462587
OD2	3.0493181966	-7.1608484863	-4.9666231056
Asp_{37B}			
CB	4.6552077353	-4.4665605591	-7.9949157093
HB1	3.6411393261	-4.6256259525	-8.3822851688
HB2	4.9575716526	-3.4794101515	-8.3624694232
CG	4.4860053371	-4.3463558849	-6.4564758476
OD1	5.0296535333	-5.1871354889	-5.7045619517
OD2	3.7651003788	-3.3709228345	-6.0520218553
Phe_{78A}			
CA	-0.6745076611	-2.2102755451	-7.7438551358
HA	-1.5599576790	-1.8613020735	-7.2019244121
C	0.5313093372	-1.7809991112	-6.9973986632
O	1.2821726073	-2.5686116397	-6.4557514898
Thr_{79A}			
N	0.7220003464	-0.4428289922	-7.1086894358
H	-0.0656523880	0.1949725804	-7.1222469040

Table S9. Cartesian coordinates of the QM atoms of the TS_{l2-l4}^{DFT} localized at B3LYP/MM level of the *substrate-assisted* phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from *E. coli*.

Dha			
C1	-2.7773028732	-1.2282716627	-3.6992131711
C2	-1.3614131631	-1.0076046224	-3.0980110964
H21	-1.2048479194	0.0804538629	-3.1829226708
H22	-1.4732980867	-1.2060829434	-2.0107608994
C3	-3.8314153539	-0.4550927239	-2.8609111247
H31	-4.8169540618	-0.6172295287	-3.3226648098
H32	-3.8570112104	-0.8717284673	-1.8422247911
O4	-2.8271865453	-0.8738967168	-5.0448971864
O5	-0.3492370766	-1.7339120525	-3.6691682033
H5	0.4021249547	-2.9960822954	-4.0044978694
O6	-3.5332590895	0.9114358902	-2.8411424016
H6	-3.0733403834	1.1413027611	-1.9615162232
ATP			
C1'	9.3027742720	-5.4961676635	2.9776549234
H1'	10.2274784091	-4.9057874436	3.0063577222
C2'	9.6625095937	-6.9882051815	2.8863315124
H2'	11.2169117795	-7.8983080454	2.2132606326
O2'	11.0680658835	-7.1066511899	2.7616340576
HT2	9.3073008405	-7.5527556100	3.7621099875
C3'	8.9330117805	-7.4629976365	1.5867206377
H3'	7.9464744336	-7.8660907962	1.8598102076
O3'	9.6853686404	-8.3735116751	0.8303420294
HT3	9.6940244017	-9.2904757842	1.2456790148
C4'	8.7576089263	-6.1484419680	0.8044469634
O4'	8.5167688988	-5.1618474553	1.8198213867
H4'	9.7009192150	-5.9365417406	0.2774769635
C5'	7.6175478749	-6.1374532928	-0.2009959766
H51	6.6544952344	-6.1268034292	0.3261667190
H52	7.6727335329	-7.0376749777	-0.8212125838
O5'	7.7476631140	-4.9670059435	-1.0148689312
PA	6.9337020347	-4.7479392138	-2.4206436613
OA1	6.4700215116	-6.0910633714	-2.9631937309
OA2	7.8440368704	-3.8780536706	-3.2468632978
OA3	5.6393459189	-3.8523451399	-1.9521021321
PB	4.0202586769	-4.3083404796	-1.8429414564
OB1	3.7916412269	-5.3306946205	-3.0133791931
OB2	3.7087130186	-4.7456122751	-0.4524296650
OB3	3.3486991839	-2.9829729518	-2.4069369696
PG	1.5328730793	-2.1789531398	-2.4841203222
OG1	0.9128791789	-2.6055898073	-1.1823957262
OG2	2.0825085649	-0.8370267476	-2.8097755461
OG3	1.3088854774	-3.3187335503	-3.6742006147

His_{56A}			
CB	-8.3380524959	-0.1688614343	-8.0634850270
HB1	-9.1553208429	0.0069545549	-7.3525248222
HB2	-7.9607097210	0.8200058627	-8.3370136432
CG	-7.2366443880	-0.9083389089	-7.3563558217
ND1	-7.0418087046	-2.2740397060	-7.3938558720
HD1	-7.6461168114	-2.9343902674	-7.8729647051
CE1	-5.9526781171	-2.5730766401	-6.6391795050
HE1	-5.5809969194	-3.5807245871	-6.5303512320
NE2	-5.4306703015	-1.4798648148	-6.0990600516
HE2	-3.7566019137	-1.0314912800	-5.3771829668
CD2	-6.2398567198	-0.4458828228	-6.5290855511
HD2	-6.0628014135	0.5698840658	-6.2143224406
His_{218A}			
CB	-2.6156662971	-6.1153640474	-5.1951416300
HB1	-1.7718261882	-5.9205882921	-5.8596092018
HB2	-2.2739714696	-6.8700085549	-4.4829713457
CG	-2.9130160171	-4.8491682309	-4.4489680168
ND1	-3.5752057174	-4.8241437438	-3.2316444639
HD1	-3.8888031295	-5.6887604089	-2.7692514517
CE1	-3.6585319506	-3.5487929265	-2.7943652560
HE1	-4.1512925666	-3.2216219231	-1.8933971344
NE2	-3.0620401632	-2.7543158748	-3.6858787459
CD2	-2.5883906474	-3.5488175260	-4.7104434458
HD2	-2.0603947970	-3.1073304542	-5.5304348824
Asp_{109A}			
CB	-1.4416605350	2.6118388820	1.3072662409
HB1	-2.0966635611	3.4712091909	1.5174248063
HB2	-1.7850726068	1.8008214762	1.9545102147
CG	-1.6890090369	2.2756798832	-0.1975621649
OD1	-2.5401305581	1.3506525597	-0.4239653689
OD2	-1.1071667884	2.9716494929	-1.0572043485
MG			
MG	4.9362697901	-6.7079179262	-4.1873262767
MG	2.9804706074	-4.1013363472	-4.4447961073
Asp_{30B}			
CB	6.9160222656	-10.0921787975	-5.5423060183
HB1	6.9106701437	-10.0137514833	-6.6325688953
HB2	7.9470046862	-9.8687611721	-5.2391224082
CG	6.0441211868	-8.9690827922	-4.9354146509
OD1	5.5199710039	-9.0607174285	-3.8047876330
OD2	5.9378087401	-7.8939999836	-5.6308588086
Asp_{35B}			
CB	0.7803158107	-7.6323415280	-5.6516943941
HB1	0.0705704674	-7.5954282290	-4.8161194865
HB2	0.2744190254	-7.1394110697	-6.4857012634
CG	1.9946665887	-6.7732072144	-5.2413782581
OD1	1.8522907607	-5.5253095449	-5.3486877827
OD2	3.0361284209	-7.3607570785	-4.8020353979

Asp_{37B}			
CB	4.6620122317	-4.4566948911	-7.9795177741
HB1	3.6340931849	-4.6400741439	-8.3159496893
HB2	4.9469728319	-3.4788534197	-8.3785224803
CG	4.5612477241	-4.3446909441	-6.4588662388
OD1	4.8647929629	-5.3308404232	-5.7354901999
OD2	4.0794225530	-3.2731846567	-5.9559483443
Phe_{78A}			
CA	-0.6187552392	-2.2449424758	-7.7640863326
HA	-1.4680194238	-1.8699221524	-7.1823152801
C	0.6340082961	-1.8140537180	-7.1140157672
O	1.4603665899	-2.5869143062	-6.6646648883
Thr_{79A}			
N	0.7765616448	-0.4656669844	-7.2224396014
H	-0.0477373064	0.1259937557	-7.1785557609

7.3. Evaluation of the effects of the mutation Glu526Lys in DHAK from *C. freundii* on the binding of poly-P

Tuning the Phosphoryl Donor Specificity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate. An Insight from Computational Studies.

Israel Sánchez-Moreno, Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía, Vicent Moliner and Eduardo García-Junceda

Published in *International Journal of Molecular Sciences* **2015**, 16, 27835-27849

International Journal of *Molecular Sciences*

MDPI

Article

Tuning the Phosphoryl Donor Specificity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate. An Insight from Computational Studies

Israel Sánchez-Moreno [1,†], Isabel Bordes [2], Raquel Castillo [2], José Javier Ruiz-Pernía [2], Vicent Moliner [2,*] and Eduardo García-Junceda [1,*]

Received: 2 October 2015; Accepted: 9 November 2015; Published: 24 November 2015
Academic Editor: Vladimír Křen

[1] Departamento de Química Bioorgánica, Instituto de Química Orgánica General, CSIC. Juan de la Cierva 3, Madrid 28006, Spain; isra.sanchez@iqog.csic.es
[2] Departament de Química Física i Analítica, Universitat Jaume I. Castellón 12071, Spain; bordes@uji.es (I.B.); rcastill@uji.es (R.C.); pernia@uji.es (J.J.R.-P.)
* Correspondences: moliner@uji.es (V.M.); eduardo.junceda@csic.es (E.G.-J.); Tel.: +34-964-728-084 (V.M.); +34-912-587-454 (E.G.-J.)
† Current address: Venter Pharma S.L., Azalea 1, Alcobendas 28109, Madrid, Spain.

Abstract: Dihydroxyacetone (DHA) kinase from *Citrobacter freundii* provides an easy entry for the preparation of DHA phosphate; a very important C3 building block in nature. To modify the phosphoryl donor specificity of this enzyme from ATP to inorganic polyphosphate (poly-P); a directed evolution program has been initiated. In the first cycle of evolution, the native enzyme was subjected to one round of error-prone PCR (EP-PCR) followed directly (without selection) by a round of DNA shuffling. Although the wild-type DHAK did not show activity with poly-P, after screening, sixteen mutant clones showed an activity with poly-phosphate as phosphoryl donor statistically significant. The most active mutant presented a single mutation (Glu526Lys) located in a flexible loop near of the active center. Interestingly, our theoretical studies, based on molecular dynamics simulations and hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) optimizations, suggest that this mutation has an effect on the binding of the poly-P favoring a more adequate position in the active center for the reaction to take place.

Keywords: biocatalysis; computational chemistry; DHAP-dependent aldolases; dihydroxyacetone kinase; enzyme directed evolution; quantum mechanics/molecular mechanics

1. Introduction

Dihydroxyacetone (DHA) kinase from *Citrobacter freundii* (*C. freundii*) provides an easy entry for the preparation of DHA phosphate, a very important C_3 building block in nature since it is used as phosphoryl donor in several enzyme-catalyzed aldol reactions [1]. Aldol addition reaction is one of the most useful tools that have the synthetic chemist for the construction of new C–C bonds [2–4]. Dihydroxyacetone phosphate (DHAP)-dependent aldolases are among the most important biocatalysts for C–C bond formation [5–10]. Their major synthetic advantage is that the stereochemistry of the two newly formed stereogenic centers is controlled by the enzymes and, moreover, the four DHAP-dependent aldolases are stereocomplementary; so, from two given substrates, it is possible to obtain the four diastereoisomers. However, it is also well known that their major drawback is their strict specificity for DHAP. Although several chemical and enzymatic routes of DHAP synthesis have been described in the literature, an efficient method of DHAP

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

preparation is still essential [11]. In this sense, our research group has developed a straightforward multi-enzyme system for one-pot C–C bond formation catalyzed by DHAP-dependent aldolases, based in the use of the recombinant ATP-dependent DHA kinase from *C. freundii* CECT 4626, for *in situ* DHAP formation [12–15]. This enzyme has the highest described catalytic efficiency for DHA phosphorylation, about one order of magnitude higher than the rest of the enzymes that have been also used for the preparation of DHAP [16]. The multi-enzyme system was completed with the *in situ* regeneration of ATP catalyzed by acetate kinase (Figure 1). Thereafter, we have reported the engineering of a new bifunctional enzyme that displays both aldolase and kinase activities in the same protein [17,18].

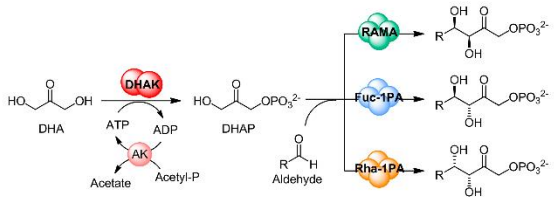


Figure 1. Multi-enzyme system for C–C bond formation catalyzed by DHAP-dependent aldolases, based in the *in situ* phosphorylation of DHA catalyzed by DHAK from *Citrobacter freundii* (*C. freundii*). DHA: Dihydroxyacetone; DHAK: dihydroxyacetone kinase; AK: acetate kinase; RAMA: rabbit muscle aldolase; DHAP: dihydroxyacetone phosphate.

Dihydroxyacetone kinase (DHAK) from *C. freundii* presents several features that make it a very interesting enzyme. It is a homodimer and each subunit is formed by two domains [19]. The K-domain is where the DHA binding site is located and in the L-domain is founded the ATP binding site (Figure 2).

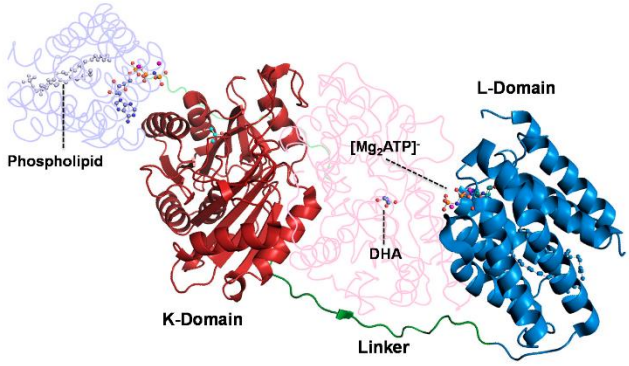


Figure 2. Crystallographic structure of DHAK from *C. freundii* strain DSM 30040 (pdb 1un9). In the N-terminal domain—K-domain—is where the DHA binding site is located. The ATP binding site is found in the C-terminal—L-domain.

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

In the dimer, the subunits are arranged in an anti-parallel way, therefore, the K-domain of one subunit is faced with the L-domain of the other subunit. The ATP binding domain is a barrel formed by eight amphipathic alpha-helix stabilized by a lipid. DHAK is the only kinase known to have an all-α nucleotide-binding domain representing a new fold group in the kinase classification scheme, as its fold is unlike any other kinase with known structure [20]. From the functional point of view, we have demonstrated that this enzyme presents a promiscuous activity [21]. Besides the phosphorylation of DHA, DHAK is able to catalyze, in the same active center, the cyclization of flavin adenine dinucleotide (FAD) to cyclic flavin mononucleotide (FMN). This catalytic promiscuity is modulated by the divalent cation that forms the complex with the phosphorylated substrate. The manganese concentration acts as a switch to turn on or off the kinase (natural) or the cyclase (promiscuous) activity.

One disadvantage of the multi-enzyme system described above is the need for ATP regeneration. To eliminate the ATP regeneration system, it is necessary to found a phosphoryl donor cheaper than ATP and, more importantly, that the final product does not inhibited the kinase. A good candidate it is the inorganic polyphosphate (poly-P). Inorganic polyphosphate (poly-P) is a linear polymer of up to hundreds of orthophosphate (P_i) residues linked by high-energy phosphoanhydride bonds [22–26]. A large amount of poly-P is routinely produced as sodium hexametaphosphate (about 13 to 18 phosphate residues) for food additives and other industrial uses, making it a phosphoryl donor inexpensive compared to ATP, acetyl-P and phosphoenolpyruvate (PEP). In addition, poly-P usually does not promote inhibition on phosphotransferases [27].

Nowadays, Computational Chemistry has emerged as a powerful tool for the analysis of reaction mechanisms in complex environments such as enzyme-catalyzed processes and the information obtained from these studies provides clues to guide the development of new and more efficient biological catalysts [28–31]. Therefore, a thorough computational based characterization of active mutants, by comparison with the wild-type enzyme, could provide information about the evolution of the substrate-protein interactions and, consequently, on the origin of the measurable activities. The results can also be used to deduce which residues of the active site are responsible for the preferential stabilization of the complex and thus, using an active mutant as starting point, to design a new catalyst capable of enhancing the rate constant of the chemical step. Herein, we describe the results obtained after a round of error-prone PCR (EP-PCR) followed (without previous selection) by a round of DNA shuffling. Theoretical calculations, based on molecular dynamics (MD) simulations and hybrid Quantum Mechanics/Molecular mechanics (QM/MM) optimizations, have been performed to study the effects of mutations on the most active mutant (Glu526Lys). A proposal of the origin of the activity will be then based on the simulations.

2. Results and Discussion

2.1. Substrate Specificity of the Wild-Type DHAK

Previously, to begin our program of Directed Evolution, we analyzed the substrate specificity of the wild-type DHAK towards the phosphoryl donor. Kinase activity of the enzyme was assayed with the five natural nucleoside triphosphates, ATP, GTP, CTP, TTP and UTP; inorganic triphosphate; and poly-P. Only ATP was substrate of the enzyme, displaying a catalytic efficiency (k_{cat}/K_M) of 6.9×10^4 $s^{-1} \cdot M^{-1}$ (K_M $3.5 \times 10^{-4} \pm 0.01$ mM and k_{cat} 24.02 ± 0.37 s^{-1}). None of the other phosphoryl donors assayed showed activity with the wild-type DHAK, at least in the limit of detection of our assay.

2.2. Generation of a DHAK Mutant Library

Given the particular structure of the DHAK (two domains connected by an 18 amino acids linker; Figure 2), we envisioned a "cassette" strategy in which it is possible to take off only the ATP-binding domain and carry out on it the evolution process, keeping the DHA-binding domain unmodified. To

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

accomplish this strategy, we took advantage of the presence in the linker region of a unique restriction site for the endonuclease *Ade*I that allow to cut-off from the pRSET-*dhak* plasmid, the l-domain where the ATP binding site is located (Figure 3).

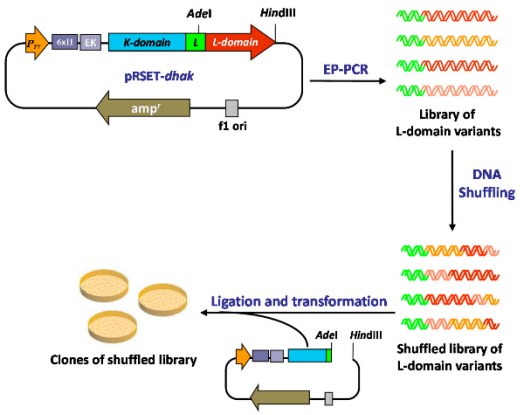


Figure 3. Directed evolution approach applied to modify the phosphoryl donor specificity of the DHAK from *C. freundii*.

To introduce genetic diversity into the sequence of the L-domain, we performed a round of EP-PCR followed directly—without previous selection—by a round of DNA shuffling. The 620 pb *domL* fragment was amplified under standard EP-PCR conditions [32–34], optimized to introduce 2–3 mutations per kilobase. The mutation rate (2.9 bases change/Kb) was verified by DNA sequencing of six randomly picked colonies of the L-domain library. The analysis of these sequence showed that DNA polymerase used in the PCR reaction (*Taq* pol) is more likely to mutate AT (67%) than GC (33%). These results were in agreement with the mutational tendency previously described in the literature [35]. The library of L-domain variants obtained in the EP-PCR step was carefully digested with DNase I, and fragments in the range of 50–100 bp were selected. The purified fragments were reassembled in a PCR reaction without primers. The number of reactions cycles (60) and the amount of DNA used as template (20% *v*/*v*, see Experimental Section) were optimized to obtain PCR products with an average size closed to full-length of the L-domain sequence (see Experimental Section).

Finally, the purified full-length L-domain sequences were amplified using standard PCR with specific primers and shuffled library of ATP-binding domains was re-introduced in the pRSET-*dhak* plasmid, restoring the whole *dhak* gene.

2.3. Screening of the DHAK Mutant Library

The mutant library was screened for their ability to phosphorylate DHA using poly-P as phosphoryl donor. Activity was measured by a coupled enzymatic assay, in which DHAP formed during the reaction in presence of poly-P was reduced to α-glycerophosphate with concomitant oxidation of NADH to NAD^+. Assay samples were prepared on 96 well plates and time-dependent decrease of NADH absorbance at 340 nm was spectrophotometrically monitored. In the primary screening, the activity was assayed in cell free extracts (CFE) from each clone of the shuffled library. Unspecific NADH oxidation was measured in reaction mixtures lacking DHA and subtracted to

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

the slope (activity) obtained in presence of DHA (Figure 4). Thus, the resulting slope should be proportional to the kinase activity using poly-P as phosphoryl donor substrate. A control population integrated by clones expressing the wild-type DHAK was analyzed following the same screening process. Both the mutant (*M*) and the control (*C*) populations followed a normal distribution according to the Kolmogorov–Smirnov test for normality [36]. The standard normal distribution of poly-P dependent activity in population *M* with respect to population *C* was obtained calculating the standard activity *Z* for each of the individuals of *M* according Equation (1).

$$Z = \frac{(x_M - \mu_M)}{\sigma_c} \quad (1)$$

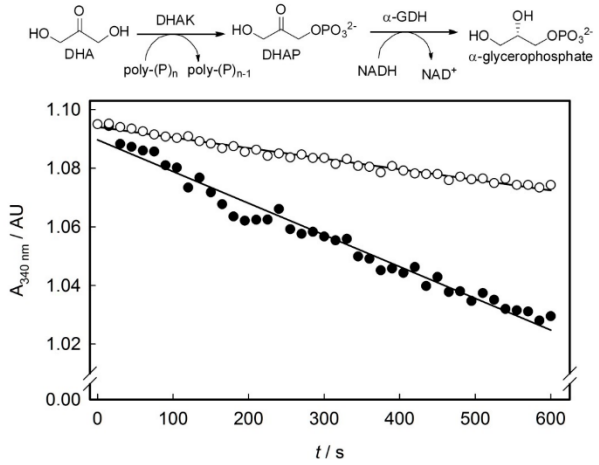


Figure 4. Enzymatic assay used to evaluate the poly-P-dependent kinase activity in the DHAK mutant library and time-dependent decrease of NADH absorbance at 340 nm corresponding to the mutant 3E4. (○) Unspecific NADH oxidation measured in absence of DHA and (●) slope obtained in presence of DHA. For details, see Section 3.6.

The statistical parameter *Z* tells us how much and in what direction—above or below—is away from the mean of the mutant population (μ_M) each individual value of slope (x_M) in the population *M* with respect to the standard deviation (σ_C) of the slope values in the control population *C*. Thus, clones with a value of $Z \geqslant 3$, had a probability ⩾98% that the slope measured was greater than the mean of the slope values in the control population. After carrying out the statistical analysis, we found 16 clones with $Z \geqslant 3$ (Figure 5). Although the statistical analyses employed returned an extremely high number of positives, we decided to express and purified the 16 positive mutants identified. The expression of the putative DHAK mutants was carried out as described in the Experimental Section and analyzed by SDS-PAGE. In all cases, a band of about 63 kDa corresponding to DHAK weight could be detected in the CFE, indicating that the mutant proteins were expressed soluble. Activities of the 16 DHAK mutants were determined after purification by affinity chromatography (IMAC). In seven of the mutants identified as positive in the statistical analyses (2D1, 3C3, 3H2, 5C1, 6A3, 6D5 and 6D6), it was impossible to detect any kinase activity

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

using poly-P as phosphoryl donor. Another group of mutants (2C1, 2H3, 4B3, 5A2, 5B1 and 6F2) was barely in the limit of detection of the employed assay. Finally, three mutants—1H2, 3E4 and 5F12—were clearly positive with polyphosphate-dependent kinase activities in the order of 10 mU per mg of protein (see Table S1). The mutant 1H2 showed the highest activity with poly-P (13.4 mU per mg of protein). This mutant presented a specific activity with ATP of 9.2 U per mg of protein, similar to the specific activity showed by the wild-type DHAK (10.7 U per mg of protein) and three orders of magnitude higher than the activity showed with poly-P.

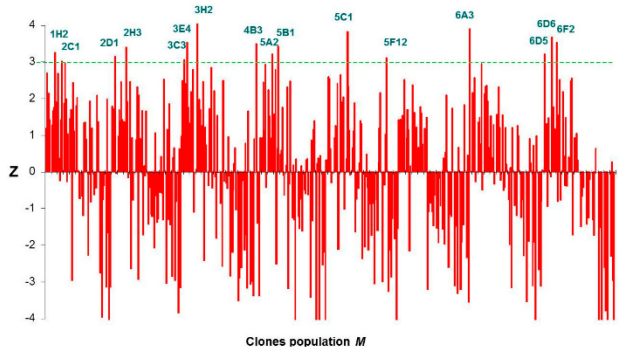


Figure 5. Graphic representation of the Z-scores of DHAK mutant population *M*.

2.4. Computational Studies of the Wild-Type and the Mutant 1H2

According to the experimental results, the computational studies were focused on the wild-type and the most active mutant: the 1H2 that presents only the mutation Glu526Lys, located on a flexible loop of the L-domain, in the surrounding of the ATP-binding site. This substitution could suggest a stronger electrostatic interaction between the positively charged lysine and the negatively charged poly-P. Nevertheless, our MD simulations suggest that mutation does not have any direct effect on the distance between the poly-P and the amino acid on position 526.

As observed in Figure 6, this distance is even shorter in the wild type than in the 1H2 (see red and blue lines, respectively). Interestingly, replacing the Glu residue on position 526 with a Lys provokes, instead, a significant approximation of the poly-P to the active site of the protein. Thus, despite both wild-type and 1H2 MD simulations starting from the same geometry, a significantly shorter distance between the closest phosphate group of poly-P and any of the Mg cations is reached after 500 ps in the 1H2 than in the wild-type. Thus, it seems that poly-P is better accommodated into the 1H2 mutant active site than in the wild type, in agreement with an experimentally observed higher catalytic activity. This effect can be shown on Figure 7, where an overlapping of both systems after the MD simulations, centered in the Mg cations, is presented. Apparently an indirect effect is produced under mutation of the protein since, despite a movement of the loop containing residue 526, no approximation is observed between the negatively charged poly-P and the new positively charged Lys 526. In fact, poly-P appears to be closer to the residue 526 in the wild-type than in the mutant, as indicated in Figure 6. This is probably due to a bending of the poly-P that favors the approximation to the active site while the new cation preferentially interacts with the solvent.

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

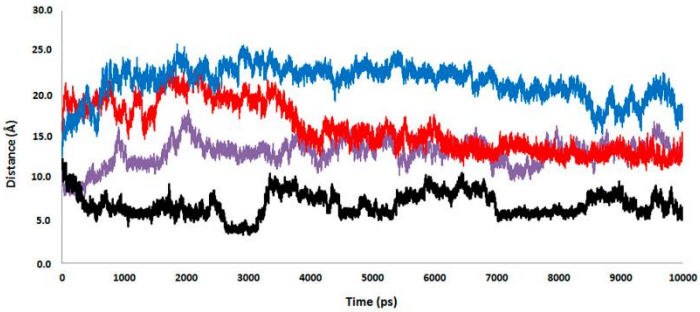


Figure 6. Time evolution of the distances between the closest atom of poly-P and the magnesium atoms in the wild-type (purple line), the magnesium atoms in the 1H2 mutant (black line), the residue 526 in the wild-type (red line), and the residue 526 in the 1H2 mutant (blue line). Results obtained from the 10 ns MD simulation.

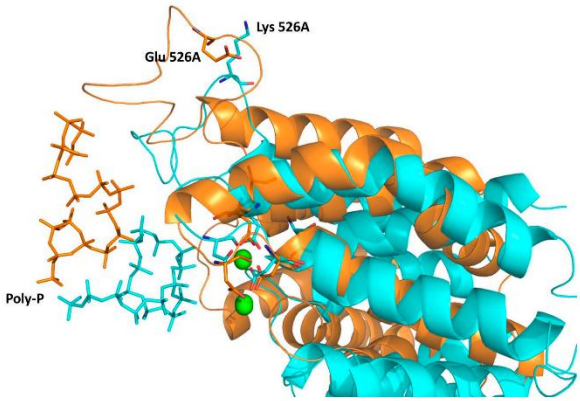


Figure 7. Overlapping of representative snapshots of the average structures in the last 500 ps of the MD simulations obtained with the wild type (gold) and 1H2 mutant (blue) complexed with the poly-P. Magnesium cations are represented as green balls. Poly-P docked in the L-domain of the 1H2 mutant (blue) and wild-type (gold) DHAK are shown as sticks. The Lys that substitutes to the Glu in position 526 in the mutant, and aspartate residues that interact with ATP in the wild-type DHAK, are also shown as sticks.

In order to confirm and quantify the interaction energy established between the poly-P and both proteins, 10 structures were selected from the MD trajectories to perform QM/MM optimizations. Keeping in mind that no covalent bond is established between the poly-P and the protein, the QM-MM interaction energy term can be directly related with the interaction between protein and poly-P. The average value of QM–MM interaction energy term of the 10 selected structures of the wild-type and mutant were -5100 ± 1100 and -7900 ± 1600 kJ·mol^{-1}, respectively. According to

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

these results, and although an expectedly large dispersion is obtained in potential energy values, the mutation has a clear effect in increasing the binding energy. The origin of the enzyme-(poly-P) interaction can be analyzed by decomposing the total interaction energy in a sum of contributions per residues. The averaged interaction energies, electrostatic plus van der Waals, of individual residues and the two magnesium cations of both systems with the poly-P is depicted in Figure 8. Bottom panel of Figure 8 shows the difference between the values obtained in the wild-type and the 1H2 mutant. Then, residues presenting positive values on this figure mean that they interact better with the mutant than with the wild-type. As observed, there are some residues, such as Lys382A, Arg475A, Lys514A and Arg207B, which present favorable interactions on both systems. On the contrary, Arg519A is stabilizing the complex only in the wild-type while the interaction of the poly-P with Mg^{2+} cations is enhanced after mutant. A graphical comparison of both systems, shown in bottom panel of Figure 8, suggests that the increase in the activity if 1H2 can be due to an enhancement of the interactions with Lys382, Arg475 and, in a dramatic way, with the magnesium cations.

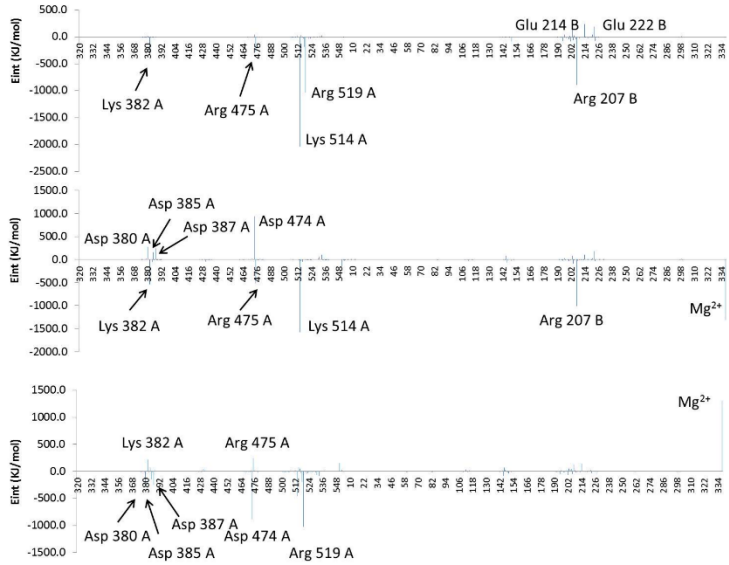


Figure 8. Contribution of individual amino acids to (poly-P)-protein interaction energy (Eint, in kJ· mol^{-1}) for the wild type (**top panel**), 1H2 mutant (**middle panel**) and differences between both systems (**bottom panel**). Residues are ordered according to the sequence, and the Mg^{2+} ions are added at the end. Results derived from the QM–MM interaction term of the QM/MM optimizations.

On the other hand, non-favorable interactions with Asp380A, Asp385A, Asp387A and Asp474A, and the loss of the originally favorable interaction observed with Arg519A in the wild-type, are observed after mutation. It is important to point out that approximation of the negatively charged poly-P to the active site is ineludible associated with a repulsive interaction with the aspartate residues (especially those anchoring the Mg^{2+} cations and, in particular Asp380A, Asp385A and Asp387A).

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

Finally, analysis of the Root Mean Square Fluctuation (RMSF) of the key residues computed from the last 5 ns of the MD simulation on the wild-type and 1H2 mutant show that the later present a greater mobility, especially on the flexible loop, what could be related with the better accommodation of the poly-P. In particular, Arg519 mobility, together with Ala520 and Ser521, significantly increases after mutation of residue 526 (Figure 9). Therefore, further efforts to improve the phosphoryl donor specificity of DHAK from ATP to an inorganic poly-P should be focused in residues around Arg519 in the protein sequence. Finally, it is important to point out that these conclusions derived from the interactions that are important for poly-P fixation should be further tested by studying the next step of the catalytic process; the chemical reaction.

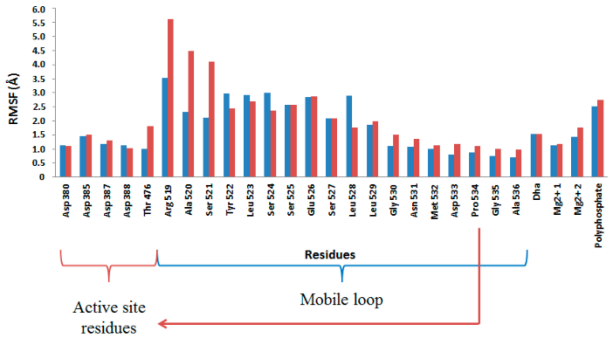


Figure 9. Root Mean Square Fluctuation (RMSF, in Å) of key residues computed from the last 5 ns of the MD simulation on the wild-type (blue bars) and 1H2 mutant (red bars).

3. Experimental Section

3.1. Materials and General Procedures

UV/Visible spectra were recorded on a Spectra Max Plus 384 spectrophotometer at 25 °C. SDS-PAGE was performed in a Mini-Protean 3 Cell Electrophoresis Unit (BioRad, Hercules, CA, USA) using 10% and 5% acrylamide in the separating and stacking gels, respectively. Gels were stained with Coomassie brilliant blue R-250 (Applichem GmBH, Darmstadt, Germany). Electrophoreses were always run under reducing conditions, in the presence of 5% β-mercaptoethanol. Protein and DNA gels were quantified by densitometry using GeneGenius Gel Documentation and Analysis System (Syngene, Cambridge, UK) equipped with the analysis software GeneSnap and GeneTools (Syngene, Cambridge, UK). α-GDH and TIM were purchased from Sigma-Aldrich (St. Louis, MO, USA). Restriction enzymes, *Taq* polymerase and T4-DNA ligase were purchased from MBI Fermentas AB (Vilnius, Lithuania). *Citrobacter freundii* CECT 4626 was provided from the Spanish Type Culture Collection (Valencia, Spain). *E. coli* BL21 (DE3) competent cells were purchased from Stratagene Co. (San Diego, CA, USA). PCR primers were purchased from Isogen Life science (Utrecht, The Netherlands) and the pRSET-A expression vector was purchased from Invitrogen Co. (Carlsbad, CA, USA). IPTG was purchased from Applichem GmBH (Darmstadt, Germany). Plasmids and PCR purification kits were from Promega (Madison, WI, USA) and DNA purification kit from agarose gels was from Eppendorf (Hamburg, Germany). Nickel-iminodiacetic acid (Ni^{2+}-IDA) agarose was supplied by Agarose Bead Technologies (Madrid, Spain). All other chemicals were purchased from commercial sources as reagent grade.

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

3.2. Error-Prone PCR Generated Library

The *dhak* gene encoding DHAK was cloned in the pRSET expression vector as previously described [21]. EP-PCR was performed on the 620 pb *dhak* fragment encoding the L-domain of DHAK using the following specific primers: leftward primer, 5′-GCGTAGCGCACGCGT-3′ (*Ade*I site underlined) and rightward primer, 5′-TTCTAAAGCTTTTAGCCCAGCTCACT-3′ (*Hind*III site underlined). The reaction mixture was prepared in a final volume of 100 µL containing 1 µM of each primer, dNTPs (0.2 mM each), 3.0 mM $MgCl_2$, 50 mM KCl, 0.2 mM $MnCl_2$, 1% DMSO, and 0.5 U *Taq* polymerase. EP-PCR was carried out using the following program: 94 °C for 2 min (1 cycle); 94 °C for 1.0 min, 55 °C for 1.0 min, 72 °C for 42 s (40 cycles); and 72 °C for 10 min (1 cycle). To verify the obtained mutation rate, a part of the L-domain variants library of was cloned in the plasmid pGEM®-T Easy and transformed in *E. coli* DH5α cells.

3.3. In Vitro Recombination by DNA Shuffling

DNA shuffling was done according to the method of Stemmer [37,38]. The pool of L-domain variants obtained by EP-PCR was digested during 5 min at RT with 0.14 U of DNase I in Tris-HCl buffer (50 mM; pH 7.4) containing $MgCl_2$ 1 mM. DNA fragments in the range of 50–100 bp were purified from the agarose gel. After that, 20% *v*/*v* of the purified DNA fragments were recombined in a PCR reaction without primers containing dNTPs (0.4 mM each), $MgCl_2$ (3.0 mM) and *Taq* polymerase (0.5 U). The thermal cycling program was: 94 °C for 2 min (1 cycle); 94 °C for 1.0 min, 55 °C for 1.0 min, 72 °C for 42 s (60 cycles); and 72 °C for 10 min (1 cycle). Finally, after 1:100 dilution of this primerless PCR product into a PCR mixture containing the specific primers described above, a single product of 620 bp corresponding to the length of the L-domain was amplified using 40 additional cycles. The resultant shuffled library of L-domain variants was ligated into the plasmid pRSET-*dhak*, previously digested with the endonucleases *Ade*I and *Hind*II. The digested plasmid contained the gene sequence encoding K-domain of the protein, which was necessary for restore the complete size of *dhak* gene in conjunction with L-domains variants. The resultant plasmids were transformed in *E. coli* BL21 (DE3) yielding over 1000 clones in a first generation.

3.4. Screening of DHAK Mutant Library

Masterplates were replicated twice into 96-well plates containing LB medium (150 µL) with ampicillin (2.5 µg/mL) and incubated at 37 °C with stirring at 200 rpm. The plates were sealed to prevent evaporation and after 90 min, protein expression was induced by addition of 0.5 mM of isopropyl-β-D-thiogalactopyranoside (IPTG). Plates were incubated in the same conditions for 14 h. Cells were harvested by centrifugation (5 min at 3000 rpm). Each cell pellet was dispersed in Tris-HCl buffer (50 mM; pH 8.0) and 10 µL of a solution containing EDTA (50 mM; pH 8.2) and lysozyme (0.35 mg/mL) were added. The plates were gently stirred at RT over 1 h and after they were kept at 4 °C overnight. Plates were centrifuged at 3000 rpm for 5 min and the cell free extracts (CFE) were transferred to new 96-well plates to carry out the activity assay. Screening was based on DHAK activity assay previously described [21] but using poly-P as phosphoryl donor instead of ATP. In this assay, DHAP formed by phosphorylation of DHA with poly-P is reduced to α-glycerophosphate with concomitant oxidation of NADH to NAD^+ and time-dependent decrease of absorbance at 340 nm due to NADH oxidation is spectrophotometricaly monitored. To carry out the assay, to each well containing 15 µL of CFE, was added 0.3 mL of reaction mixture containing Tris-HCl (40 mM; pH 8.0), DHA (2.5 µM), NADH (0.2 µM), poly-P (150 µM), $MgSO_4$ (2.5 µM) and αGDH (2.6 U).

3.5. Expression and Purification of Selected Mutants

Expression of wild-type DHAK and the 16 positive mutants was performed as previously described [21]. Expression was induced by IPTG addition (1 mM) when the culture reached an $O.D_{600nm}$ of 0.5–0.6. After induction the culture temperature was dropped to 30 °C. The culture

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

was maintained O/N and then centrifuged at 10,000× *g* during 10 min at 4 °C. The resulting pellet was treated with lysozyme and DNase for protein extraction. Recombinant proteins containing an N-terminal 6xHis tag was purified in a Ni^{+2}-IDA-agarose column pre-equilibrated with sodium phosphate buffer (20 mM; pH 7.5). Protein elution was achieved with the same buffer containing imidazole 0.25 M. Mutant 1H2 was further purified by size-exclusion chromatography on a HiLoad 26/60 Superdex 75 PG column controlled using the AKTA-FPLC system (GE Healthcare Life Science, Little Chalfont, UK). Purification was carried out in 50 mM phosphate buffer pH 7.2 containing NaCl (0.15 M) at a constant flow rate of 1.0 mL/min (Figure 10).

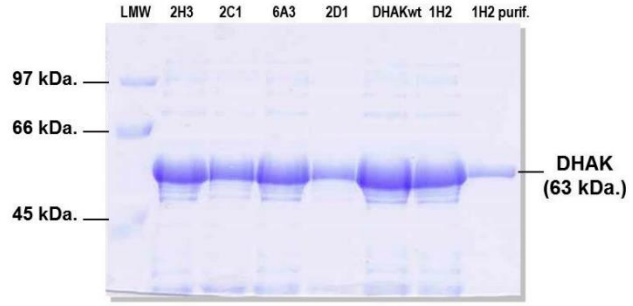


Figure 10. SDS-PAGE analysis of the expression of five mutants and the wild-type DHAK (DHAKwt). Left lane shows mutant 1H2 after purification.

3.6. Enzyme Activity Assays of Purified Mutants and wtDHAK

Phosphorylation of DHA was measured spectrophotometrically in a coupled enzymatic assay based in the reduction of DHAP to α-glycerophosphate, catalyzed by αGDH with concomitant oxidation of NADH to NAD+. The assays were run at 25 or 37 °C following the decrease of absorbance at 340 nm (ε^{340}_{NADH} = 6220 $cm^{-1} \cdot M^{-1}$) for 10 min in reaction mixtures of 1 mL containing Tris–HCl (40 mM, pH 8.0), DHA (2.5 µmol), NADH (0.2 µmol), ATP (5.0 µmol), $MgSO_4$ (5.0 µmol), αGDH (1.26 U), TIM (12.6 U) and DHAK. When different phosphoryl donors were assayed, ATP was substituted by GTP, CTP, TTP, UTP, inorganic triphosphate or poly-P. Control activity assays were carried out with every mutant enzyme where the phosphoryl donor or DHA were not initially included. Also, blanks without enzyme were performed to be sure that no activity is observed when the enzyme is absent. After addition of the corresponding substrate, differences in the slopes were measured (Figure S1). One unit of kinase activity was defined as the amount that produces 1 µmol of DHAP per min under the above conditions.

3.7. Computational Model and Methods

The first step of our protocol was setting up the system for performing the MD simulations. The initial coordinates of the protein and the phospholipid were taken from the X-ray structure of the apo form of DHAK from *C. freundii* strain DSM 30040 (pdb 1UN8) [19]. Some missing residues located on the flexible loop of the L-domain were manually added within the help of Molden program [39]. In particular, and since the PDB file contained the full sequence of the protein, the missed residues were introduced in Molden together the 15 upstream and downstream residues in order to obtain the most favorable conformation from the missing loop. Anyway, as explained below, MD simulations will be run to explore different conformation of this flexible loop. The coordinates of DHA and magnesium cations were taken from the PDB file 1UN9 that corresponds to the Dha/ANP form [19]. Later, since

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

the standard pKa values of ionizable groups can be shifted by local protein environments, an accurate assignment of the protonation states of all these residues at pH = 7 was carried out. Recalculation of the pKa values of the titratable amino acids has been done using the empirical PropKa program of Jensen *et al.* [40,41]. According to these results, all residues were found at their standard protonation state in aqueous solution, except His404 that was double protonated. Then, a 16 monomers poly-P was docked into the active site of the protein by placing a phosphate group in a position equivalent to the one the ATP would present for the chemical reaction to take place. Afterwards, in order to avoid overlapping between the poly-P and the protein, the former was displaced following a perpendicular vector. Then the mutated enzyme was generated by replacing Glu526 by a Lys. The 30 and 28 counter ions (Na^+) were placed into optimal electrostatic positions around the wild type protein and mutant, respectively, in order to obtain electro neutrality. The systems were placed in a box of pre-equilibrated waters (100 Å × 80 Å × 80 Å), using the center of mass of the poly-P as the geometrical center. Any water with an oxygen atom lying in a radius of 2.8 Å from a heavy atom of the protein or counter ions was deleted. Afterwards, series of optimization algorithms (steepest descent conjugated gradient and L-BFGS-B) were applied. The two systems were equilibrated by means of 100 ps of Langevin–Verlet MD (NVT) [42] at temperature of 300 K using the fDYNAMO library [43]. Finally, 10 ns of MD simulations were performed within the NAMD parallel molecular dynamics code [44]. During the MD simulations, the atoms of the protein and poly-P were described using the OPLS-AA [45,46] force field while TIP3P force field [47] was used for the water molecules. Cutoffs for the nonbonding interactions were applied using a switching function, within a radius range from 14.0 to 16.0 Å, employing in all the simulations periodic boundary conditions. The analysis of the total energy and the RMSD of the wild-type and the mutant shows that both systems can be considered as equilibrated after the MD simulations (see Figures S2 and S3). Moreover, the DHA appears to be well fitted in the active site, with a pattern of interactions with the residues of the active site that was already observed in the initial X-ray diffraction structure (see Figure S4).

After the MD simulations, ten different structures from the last 5 ns of the two MD trajectories were selected to perform full hybrid QM/MM optimizations with fDYNAMO library. Treating the poly-P quantum mechanically and the protein molecular mechanically has the advantage of the inclusion of quantum effects such as poly-P polarization upon binding. Moreover, since no covalent bonds are established between the poly-P and the protein, interaction energy (that can be decomposed by residue) fits with the QM–MM interaction energy term of the full Hamiltonian and, consequently, the reported interaction energies include electrostatic and van der Waals terms computed quantum mechanically. Moreover, as the largest part of the system is described classically, enough sampling can be obtained at reasonable computational cost. All in all, a deeper understanding of the binding process can be obtained from more reliable results. In these calculations, the poly-P was treated by the semi-empirical AM1 [48] Hamiltonian, while the protein and water molecules were treated by OPLS-AA and TIP3P force fields, respectively.

4. Conclusions

Using an approach of Enzyme Directed Evolution, we have been able to generate a new activity on the DHAK from *C. freundii*. After a round of error-prone PCR followed by a round of artificial recombination by DNA shuffling, it was possible to detect a mutant showing kinase activity using poly-P as phosphate group donor in the order of tens of mU. This mutant, termed 1H2, presents only one mutation consisting of the change of Glu526, located in an unstructured loop in the surrounding of the ATP binding site, to a Lys. Computational studies allow concluding that this mutation does not increase the interaction of the new positively charged amino acid with the negatively charged poly-P. Instead, an indirect effect is generated that promotes a displacement of the loop that pushes the L-domain α -helix allowing the poly-P coming closer to the ATP binding site and then favoring the reaction to take place.

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/16/11/26073/s1.

Acknowledgments: We thank Rocío Martín-Hoyos for her contribution to the optimization of the conditions to perform the directed evolution program. This work was supported by grants from the Spanish Ministerio de Ciencia e Innovación (Grants CTQ2007-67403/BQU, CTQ2010-15418 and PI11/01436), Ministerio de Economía y Competitividad (project CTQ2012-36253-C03-01), Comunidad de Madrid (Grant S2009/PPQ-1752), Generalitat Valenciana (PROMETEOII/2014/022), and Universitat Jaume I (project P1 1B2014-26). Israel Sánchez-Moreno was a predoctoral fellow from Comunidad de Madrid.

Author Contributions: Israel Sánchez-Moreno performed the directed evolution program, the biochemical studies and wrote the manuscript; Isabel Bordes performed calculations and prepared figures; Raquel Castillo and José Javier Ruiz-Pernía supervised the calculations and analyzed the data; Vicent Moliner and Eduardo García-Junceda conceived and designed the research, analyzed the data and wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Enders, D.; Voith, M.; Lenzen, A. The dihydroxyacetone unit—A versatile C3 building block in organic synthesis. *Angew. Chem. Int. Ed.* **2005**, *44*, 1304–1325. [CrossRef] [PubMed]
2. Li, C.J. Organic reactions in aqueous media with a focus on carbon-carbon bond formations: A decade update. *Chem. Rev.* **2005**, *105*, 3095–3166. [CrossRef] [PubMed]
3. Palomo, C.; Oiarbide, M.; García, J.M. Current progress in the asymmetric aldol addition reaction. *Chem. Soc. Rev.* **2004**, *33*, 65–75. [CrossRef] [PubMed]
4. Alcaide, B.; Almendros, P. The direct catalytic asymmetric cross-aldol reaction of aldehydes. *Angew. Chem. Int. Ed.* **2003**, *42*, 858–860. [CrossRef] [PubMed]
5. Clapés, P.; Joglar, J. Enzyme-catalyzed aldol additions. In *Modern Methods in Stereoselective Aldol Reactions*; Mahrwald, R., Ed.; Wiley-VCH: Weinheim, Germany, 2013; pp. 475–527.
6. Fesko, K.; Gruber-Khadjawi, M. Biocatalytic methods for C–C bond formation. *ChemCatChem* **2013**, *5*, 1248–1272. [CrossRef]
7. Müller, M. Recent developments in enzymatic asymmetric C–C bond formation. *Adv. Synth. Catal.* **2012**, *354*, 3161–3174. [CrossRef]
8. Brovetto, M.; Gamenara, D.; Saenz Méndez, P.; Seoane, G.A. C–C bond-forming lyases in organic synthesis. *Chem. Rev.* **2011**, *111*, 4346–4403. [CrossRef] [PubMed]
9. Clapés, P.; Fessner, W.-D.; Sprenger, G.A.; Samland, A.K. Recent progress in stereoselective synthesis with aldolases. *Curr. Opin. Chem. Biol.* **2010**, *14*, 154–167. [CrossRef] [PubMed]
10. Iturrate, L.; García-Junceda, E. DHAP-dependent aldolases in the core of multistep processes. In *Multi-Step Enzyme Catalysis: Biotransformations and Chemoenzymatic Synthesis*; García-Junceda, E., Ed.; Wiley-VCH: Weinheim, Germany, 2008; pp. 61–81.
11. Schümperli, M.; Pellaux, R.; Panke, S. Chemical and enzymatic routes to dihydroxyacetone phosphate. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 33–45. [CrossRef] [PubMed]
12. Sánchez-Moreno, I.; García-García, J.F.; Bastida, A.; García-Junceda, E. Multienzyme system for dihydroxyacetone phosphate-dependent aldolase catalyzed C–C bond formation from dihydroxyacetone. *Chem. Commun.* **2004**, *14*, 1634–1635.
13. Sánchez-Moreno, I.; Iturrate, L.; Doyagüez, E.G.; Martínez, J.A.; Fernández-Mayoralas, A.; García-Junceda, E. Activated α,β-unsaturated aldehydes as substrate of dihydroxyacetone phosphate (DHAP)-dependent aldolases in the context of a multienzyme system. *Adv. Synth. Catal.* **2009**, *351*, 2967–2975. [CrossRef] [PubMed]
14. Camps Bres, F.; Guérard-Hélaine, C.; Hélaine, V.; Fernandes, C.; Sánchez-Moreno, I.; Traïkia, M.; García-Junceda, E.; Lemaire, M. L-Rhamnulose-1-phosphate and L-fuculose-1-phosphate aldolase mediated multi-enzyme cascade systems for nitrocyclitol synthesis. *J. Mol. Catal. B: Enzym.* **2015**, *114*, 50–57. [CrossRef]
15. Oroz-Guinea, I.; Hernández, K.; Camps Bres, F.; Guérard-Hélaine, C.; Lemaire, M.; Clapés, P.; García-Junceda, E. L-Rhamnulose-1-phosphate aldolase from Thermotoga maritima in organic synthesis: One-pot multistep reactions for the preparation of imino- and nitrocyclitols. *Adv. Synth. Catal.* **2015**, *357*, 1951–1960. [CrossRef]

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

16. Itoh, N.; Tujibata, Y.; Liu, J.Q. Cloning and overexpression in *Escherichia coli* of the gene encoding dihydroxyacetone kinase isoenzyme I from Schizosaccharomyces pombe, and its application to dihydroxyacetone phosphate production. *Appl. Microbiol. Biotechnol.* **1999**, *51*, 193–200. [CrossRef] [PubMed]
17. Iturrate, L.; Sanchez-Moreno, I.; Doyaguez, E.G.; Garcia-Junceda, E. Substrate channelling in an engineered bifunctional aldolase/kinase enzyme confers catalytic advantage for C–C bond formation. *Chem. Commun.* **2009**, *13*, 1721–1723. [CrossRef] [PubMed]
18. Iturrate, L.; Sánchez-Moreno, I.; Oroz-Guinea, I.; Pérez-Gil, J.; García-Junceda, E. Preparation and characterization of a bifunctional aldolase/kinase enzyme: A more efficient biocatalyst for C–C bond formation. *Chem. Eur. J.* **2010**, *16*, 4018–4030. [CrossRef] [PubMed]
19. Siebold, C.; Arnold, I.; García-Alles, L.F.; Baumann, U.; Erni, B. Crystal structure of the Citrobacter freundii dihydroxyacetone kinase reveals an eight-stranded α-helical barrel ATP-binding domain. *J. Biol. Chem.* **2003**, *278*, 48236–48244. [CrossRef] [PubMed]
20. Cheek, S.; Ginalski, K.; Zhang, H.; Grishin, N.V. A comprehensive update of the sequence and structure classification of kinases. *BMC Struct. Biol.* **2005**, *5*. [CrossRef] [PubMed]
21. Sánchez-Moreno, I.; Iturrate, L.; Martín-Hoyos, R.; Jimeno, M.L.; Mena, M.; Bastida, A.; García-Junceda, E. From Kinase to Cyclase: An unusual example of catalytic promiscuity modulated by metal switching. *ChemBioChem* **2009**, *10*, 225–229. [CrossRef] [PubMed]
22. Kornberg, A. Inorganic polyphosphate: Toward making a forgotten polymer unforgettable. *J. Bacteriol.* **1995**, *177*, 491–496. [PubMed]
23. Kornberg, A.; Rao, N.N.; Ault-Riche, D. Inorganic polyphosphate: A molecule of many functions. *Annu. Rev. Biochem.* **1999**, *68*, 89–125. [CrossRef] [PubMed]
24. Brown, M.R.; Kornberg, A. Inorganic polyphosphate in the origin and survival of species. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16085–16087. [CrossRef] [PubMed]
25. Brown, M.R.; Kornberg, A. The long and short of it—polyphosphate, PPK and bacterial survival. *Trends Biochem. Sci.* **2008**, *33*, 284–290. [CrossRef] [PubMed]
26. Rao, N.N.; Gómez-García, M.R.; Kornberg, A. Inorganic polyphosphate: Essential for growth and survival. *Annu. Rev. Biochem.* **2009**, *78*, 605–647. [CrossRef] [PubMed]
27. Haeusler, P.A.; Dieter, L.; Rittle, K.J.; Shepler, L.S.; Paszkowskiy, A.L.; Moe, O.A. Catalytic properties of Escherichia coli polyphosphate kinase: An enzyme for ATP regeneration. *Biotechnol. Appl. Biochem.* **1992**, *15*, 125–133. [PubMed]
28. Ferrer, S.; Martí, S.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertrán, J. Molecular mechanism of promiscuous MbtI. Improving the secondary activity by a computational design. *Theor. Chem. Acc.* **2011**, *128*, 601–607. [CrossRef]
29. Senn, H.M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229. [CrossRef] [PubMed]
30. Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Computer-aided rational design of catalytic antibodies: The 1F7 case. *Angew. Chem. Int. Ed.* **2007**, *46*, 286–290. [CrossRef] [PubMed]
31. Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Predicting an improvement of secondary catalytic activity of promiscuous isochorismate pyruvate lyase by computational design. *J. Am. Chem. Soc.* **2008**, *130*, 2894–2895. [CrossRef] [PubMed]
32. Leung, D.W.; Chen, E.; Goeddel, D.V. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1989**, *1*, 11–15.
33. Cirino, P.C.; Mayer, K.M.; Umeno, D. Generating mutant libraries using error-prone PCR. In *Methods in Molecular Biology, Volume 231: Directed Evolution Library Creation*; Arnold, F.H., Georgiou, G., Eds.; Humana Press Inc.: Totowa, NJ, USA, 2003; pp. 3–9.
34. McCullum, E.O.; Williams, B.A.R.; Zhang, J.; Chaput, J.C. Random mutagenesis by error-prone PCR. In *Methods in Molecular Biology, Volume 634: In Vitro Mutagenesis Protocols*, 3rd ed.; Braman, J., Ed.; Humana Press Inc.: Totowa, NJ, USA, 2010; pp. 103–109.
35. Rasila, T.S.; Pajunen, M.I.; Savilahti, H. Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, Escherichia coli mutator strain, and hydroxylamine treatment. *Anal. Biochem.* **2009**, *388*, 71–80. [CrossRef] [PubMed]

Int. J. Mol. Sci. **2015**, *16*, 27835–27849

36. Kvam, P.H.; Vidakovic, B. *Nonparametric Statistics with Applications to Science and Engineering*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007.
37. Stemmer, W.P.C. DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747–10751. [CrossRef] [PubMed]
38. Stemmer, W.P.C. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **1994**, *370*, 389–391. [CrossRef] [PubMed]
39. Schaftenaar, G.; Noordik, J.H. Molden: A pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123–134. [CrossRef] [PubMed]
40. Bas, D.C.; Rogers, D.M.; Jensen, J.H. Very fast prediction and rationalization of pK_a values for protein-ligand complexes. *Proteins: Struct. Funct. Bioinf.* **2008**, *73*, 765–783. [CrossRef] [PubMed]
41. Li, H.; Robertson, A.D.; Jensen, J.H. Very fast empirical prediction and rationalization of protein pK_a values. *Proteins: Struct. Funct. Bioinf.* **2005**, *61*, 704–721. [CrossRef] [PubMed]
42. Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98–103. [CrossRef]
43. Field, M.J. *A Practical Introduction to the Simulation of Molecular Systems*; Cambridge University Press: Cambridge, UK, 2007.
44. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]
45. Jorgensen, W.L.; Tiradorives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666. [CrossRef]
46. Pranata, J.; Wierschke, S.G.; Jorgensen, W.L. OPLS potential functions for nucleotide bases. Relative association constants of hydrogen-bonded base pairs in chloroform. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819. [CrossRef]
47. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
48. Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. Development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909. [CrossRef]

Supplementary Materials: Tuning the Phosphoryl Donor Specificity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate. An Insight from Computational Studies

Israel Sánchez-Moreno, Isabel Bordes, Rocío Martín-Hoyos, Raquel Castillo, José Javier Ruiz-Pernia, Vicent Moliner and Eduardo García-Junceda

Table S1. Polyphosphate-dependent kinase activities of the wild-type and mutant enzymes [a].

Enzyme	wtDHAK	1H2	2C1	2D1	2H3	3C3	3E4	3H2	4B3
Specific activity ($mU \cdot mg^{-1}$)	n.d [b]	13.4	2.6	n.d	1.9	n.d	7.7	n.d	0.1
Enzyme	-	**5A2**	**5B1**	**5C1**	**5F12**	**6A3**	**6D5**	**6D6**	**6F2**
Specific activity ($mU \cdot mg^{-1}$)	-	3.0	2.7	n.d	10.5	n.d	n.d	n.d	0.4

[a] Nomenclature of mutant enzymes is arbitrary and corresponds to the coordinates in the micro-plate; [b] n.d = not detected.

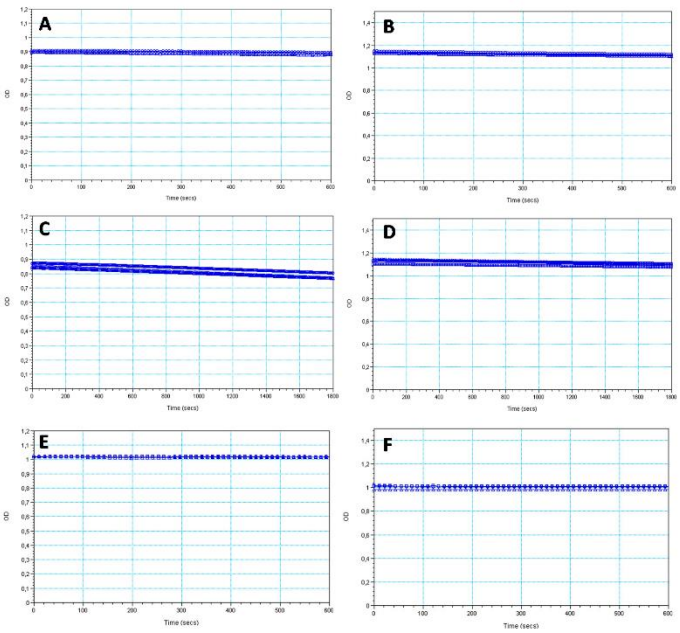


Figure S1. Absorbance at 340 nm was recorded along the time in triplicate reactions containing poly-P and mutant 1H2 (**A**); The same assay was performed with wild type DHAK enzyme (**C**); When DHA was added into the pre-incubated reactions, significant slope differences were observed in 1H2 mutant reactions (**B**) but no in DHAK wt ones (**D**); In order to exclude any spontaneous chemical phosphorylation, control reactions were incubated in absence of the corresponding enzyme: poly-P/DHA mixture without 1H2 mutant (**E**);and ATP/DHA mixture without DHAK wt (**F**).

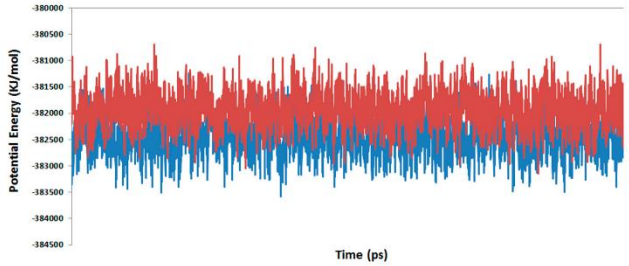


Figure S2. Time evolution of the potential energy of the full system in the wild-type (blue line) and mutated enzyme (red line).

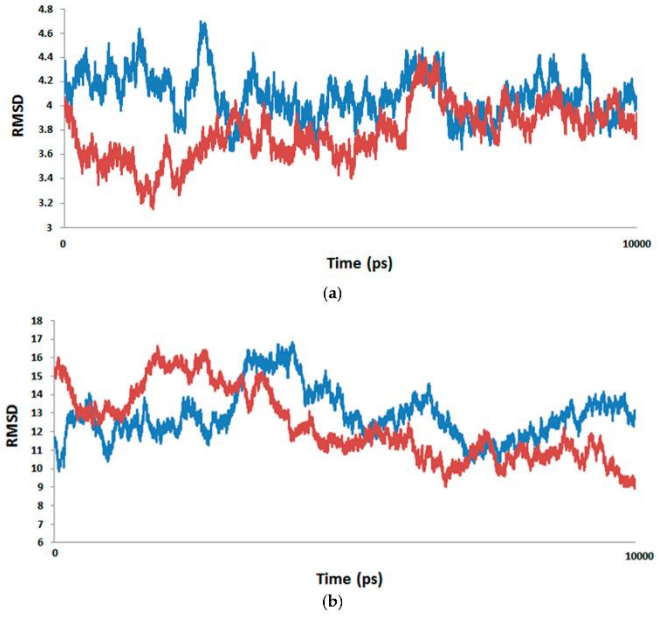


Figure S3. (**a**) RMSD of the carbon-alpha atoms of the protein in the wild type (blue line) and in the mutated enzyme (red line); (**b**) RMSD of the phosphorus atoms of poly-P along the last 5 ns of MD simulation in the wild type (blue line) and mutated enzyme (red line).

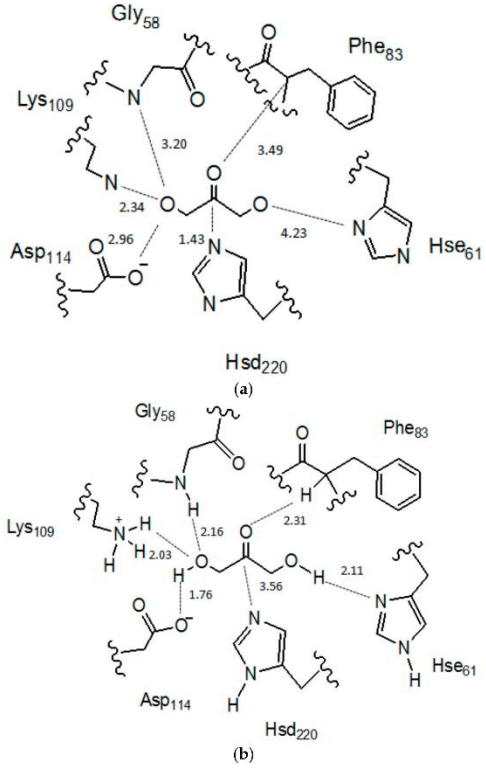


Figure S4. Schematic representation of the active site in the initial X-ray diffraction structure 1UN9 (**a**); and in the wild-type after 10 ns of MD simulations (**b**). Key distances between DHA and the residues of the active site are reported in Å.

7.4. Study of the phosphoryl transfer reaction from inorganic poly-P to Dha in DHAK from *C. freundii*

A Computational Study of the Phosphoryl Donor Activity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate.

Isabel Bordes, Raquel Castillo and Vicent Moliner

Submitted for publication to *International Journal of Quantum Chemistry*

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

Computational Study of the Phosphoryl Donor Activity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate

Isabel Bordes Pastor,[1] Eduardo García-Junceda,[2] Raquel Castillo Solsona, [1] Vicent Moliner Ibáñez[1]

Correspondence to:

Raquel Castillo Solsona (E-mail: rcastill@uji.es) tel: +34964728097

Vicent Moliner Ibáñez (E-mail: moliner@uji.es) tel: +34964728084

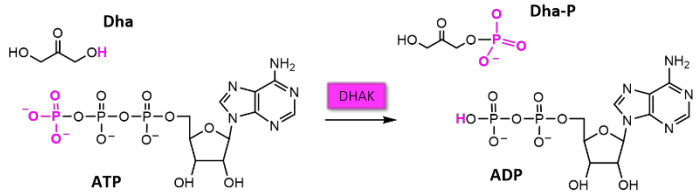
[1] *Isabel Bordes Pastor, Raquel Castillo Solsona, Vicent Moliner Ibáñez*
Departament de Química Física i Analítica, Universitat Jaume I, Castelló, Spain, 12071
[2] *Eduardo García-Junceda*
Departamento de Química Bioorgánica, Instituto de Química Orgánica General, CSIC, Juan de la Cierva 3, Madrid, Spain, 28006

ABSTRACT

Adenosine triphosphate (ATP) is the main biological phosphoryl donor required in many enzymes including dihydroxyacetone kinases (DHAKs) that convert dihydroxyacetone (Dha) into dihydroxyacetone phosphate (Dha-P), a key species with potential applications in synthesis. Herein it is presented a theoretical study of the molecular mechanism for the phosphoryl transfer reaction from an inorganic polyphosphate to Dha catalyzed by DHAK from *C.freundii*. This is part of a project devoted to modify the phosphoryl donor specificity of this enzyme avoiding the use of the problematic direct addition of ATP. Based on the use of hybrid QM/MM potentials, with the QM region described by semiempirical and DFT methods, the reaction mechanism of the wild type enzyme and the most active experimentally measured mutant (Glu526Lys) with poly-P as phosphoryl donor has been explored to elucidate the origin of the activity of this mutant. The similar energy barriers obtained in both systems confirm our previous studies on the binding step (*Int. J. Mol. Sci.* **2015**, *16*, 27835-27849) suggesting that this mutation favours a more adequate position of the poly-P in the active site for the following step, the chemical reaction, to take place.

Introduction

Phosphorylation, the chemical process that implies the transfer of the phosphoryl group from a phosphate ester or anhydride to a nucleophile,[1] is essential in many procedures occurring in cells of living organisms such as propagation, signal transduction and replication of genetic material.[2-4] In principle, phosphorylation reactions can occur through two likely mechanisms:[3,5-7] associative, where nucleophilic attack is produced before departure of living group, and dissociative, where living group departure occurs previously to the nucleophilic attack. These reactions, in turn, may take place in a stepwise manner or by means of a single step or concerted mechanism. Adenosine triphosphate (ATP) is the main biological phosphoryl donor required in many enzymatic reactions. In fact, one of the largest protein family of enzymes, namely kinases,[8,9] catalyses the transfer of the terminal phosphate group from ATP to a substrate or other proteins[10]. Different research works have been published about protein kinases where the phosphate from an ATP molecule is transferred mainly to a serine, threonine or a tyrosine residue.[11-14] From the mechanistic point of view, some authors describe phosphorylation reactions where a conserved aspartate residue acts as a base activating the acceptor residue,[15-19] while others consider that this activation is caused by the ATP substrate itself.[20,21] The former process is categorized as *asp-assisted mechanism* and the latter is called *substrate-assisted mechanism*. An example for the first mechanism was proposed by Shi et. al,[22] for the phosphoryl transfer reaction from ATP molecule to dihydroxyacetone (Dha) substrate in dihydroxyacetone kinase (DHAK) from *Escherichia Coli* (*E.coli*). Based on a crystal structure analysis, combined mutagenesis and enzymatic activity studies, the authors claimed that the aspartate residue plays the role of a basis taking the proton from Dha. However, De Vivo et al., based on theoretical results derived from gas-phase DFT calculations, classical molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) Car-Parrinello simulations, suggested a *substrate-assisted mechanism* for the phosphorylation reaction in cyclin-dependent kinase (CDK2) where ATP was shown to take the proton from a serine residue.[21] We have recently studied the phosphoryl transfer mechanism from ATP to Dha in DHAK from *E.coli* finding the *substrate-assisted mechanism* kinetically more favorable than the *asp-assisted mechanism*.[23]



Scheme 1. Schematic representation of the reaction between ATP and Dha generating Dha-P catalysed by DHAK.

DHAKs are divided into two classes, those using ATP as the phosphoryl donor in animals, plants and some bacteria, and those using the phosphoenol pyruvate carbohydrate (phosphotransferase system, PTS) to provide the phosphoryl group as observed in most bacteria.[22,24-29] DHAKs phosphorylate Dha converting it into dihydroxyacetone phosphate

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

(Dha-P), as shown in Scheme 1. Dha-P is a very important intermediate in nature since it is used as phosphoryl donor in several enzyme-catalyzed aldol reactions by Dha-P dependent aldolases.[30-32] Aldolases have been recognized as an indispensable tool for the organic synthesis due to their efficiency to form C-C bonds. These reactions occur with the formation of two new stereocenters with a stereochemistry controlled by enzymes. Then, from two given substrates, it is possible to obtain the four diastereoisomers using four different aldolases which is a powerful synthetic advantage.[33-37] Nevertheless, the major disadvantage of Dha-P dependent aldolases is their strict specificity for the expensive and unstable Dha-P.[33-35,37] This molecule is overpriced to be used in large-scale synthesis and is labile at neutral and basic pH values causing the decrease of its effective concentration with time in the enzymatic reaction media.[34] Therefore, an efficient method of Dha-P preparation is still necessary.[31,35] In this sense, García-Junceda and co-workers, on the basis of the recombinant ATP-dependent DHAK from *C. freundii*, have elaborated a straightforward multi-enzyme system for one-pot C–C bond formation catalyzed by DhaP-dependent aldolases for in situ Dha-P formation. In this system the ATP is regenerated *in situ* by catalysis of acetate kinase.[31,32,34,38] The fusion protein keep up both kinase and aldolase activity with a very high catalytic efficiency. Nevertheless, despite of the benefits of this multi-enzyme process, it still requires an ATP regeneration system since the direct addition of ATP is often problematic due to the formation of inhibitory products such as ADP or AMP.[39,40] ATP-regenerating systems in the reactions can be produced through the employment of biological agents including whole cells, organelles, and enzymes.[39-42] However, as mentioned, there are difficulties with these systems regarding the elevated prices of chemicals and the inaccessibility to a method for regenerating ATP from AMP.[41]

Therefore, in order to eliminate the ATP regeneration system, it is essential to found a phosphoryl donor cheaper than ATP and whose final products do not inhibit the kinase activity. A very suitable candidate is the inorganic polyphosphate (poly-P).[43,44] Inorganic poly-P is a linear polymer of up to hundreds of orthophosphate (Pi) linked by high-energy phosphoanhydride bonds (see Figure 1).[43,45]

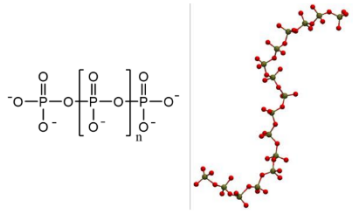


Figure 1. Chemical structure of inorganic polyphosphate polymer (left) and a ball and stick representation of a polyphosphate of n=14 (right).

The economic saving of the use of poly-P as phosphate donor is obvious. For instance, a commercial form of poly-P costing ca. $9/lb can provide ATP equivalents that would cost over $2,000/lb separately, while other phosphagens capable of regenerating ATP, such as phosphoenolpyruvate (PEP) and phosphocreatine, cost more than ATP[39]. In addition, a huge quantity of poly-P is regularly produced as sodium hexametaphosphate (about 13 to 18 residues) for industrial uses such as food additives which also makes poly-P inexpensive compared to the other phosphoryl donors.[31]

The wild type DHAK from *C. freundii* does not show activity with poly-P but sixteen mutant clones were found to exhibit certain activity with poly-P as phosphoryl donor statistically relevant.[31] The most active mutant was based on a single mutation Glu526Lys (E526K), which implies the change of a negative charge for a positive one. This mutation is located in a flexible loop close to the active site.[31] Calculations based on molecular dynamics simulations and hybrid QM/MM optimizations

of the enzyme-poly-P complex carried out in our laboratory indicate that the Lys526 could interact with poly-P stabilizing its binding with the enzyme and contributing to its correct disposition in the active site.[31]

This paper is focused on the chemical step for the wild type enzyme and the E526K mutant in order to elucidate the reaction mechanism and to compare the activity in both systems. We present a theoretical study of the molecular mechanism for the phosphoryl transfer reaction from poly-P to Dha catalyzed by DHAK from *C.freundii*. The results are based on molecular dynamics (MD) simulations and the exploration of potential energy surfaces (PES) using hybrid QM/MM potentials. The results will make it possible to check whether the activity of the mutant with poly-P as phosphoryl donor is due merely to the more reactive enzyme-poly-P binary complex,[31] or the mutation also implies a reduction on the energy barrier of the chemical step.

Computational methods

DHAK from *Citrobacter Freundii* (*C. freundii*) is an ATP-dependent DHAK consisting of a homodimer and each subunit is formed by two domains.[25,29,46,47] The Dha binding site is located in the DhaK-domain while the ATP binding site is in the DhaL-domain.

In the dimer, the subunits are disposed in an anti-parallel way. Therefore, the DhaK-domain of one subunit is faced with the DhaL-domain of the other subunit. The ATP binding domain is a barrel composed by eight amphipathic alpha-helix stabilized by a lipid (see Figure 2).[31] The phosphate groups of the nucleotide are coordinated via two magnesium ions to the side-chain carboxyl groups of aspartates.[46]

The initial coordinates of the protein and the phospholipid were taken from the X-ray structure of the apo form of DHAK from *Citrobacter Freundii* (*C. freundii*), with pdb entry 1UN8.[46] The crystal structure contains two protein chains defined as chain A and chain B. Since the full structure is symmetric, a fragment of each chain was removed obtaining a two close domain structure where the chain A fragment corresponds to the DhaL-domain, and the Chain B to the DhaK-domain. Missing residues of the flexible loop of the L-domain were manually added within the help of Molden program.[48] The coordinates of Dha and magnesium cations were taken from the PDB file 1UN9 that corresponds to the Dha/ANP form.[46]

Considering that the standard pKa values of ionizable groups can be shifted by local protein environments, an accurate assignment of the protonation state of all these residues at pH=7 was carried out.

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

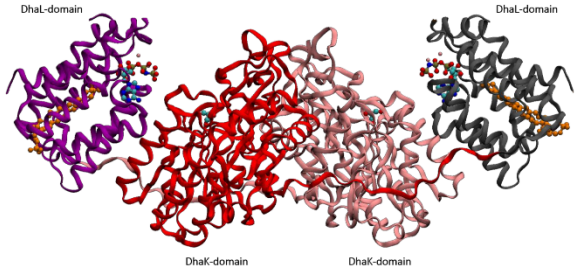
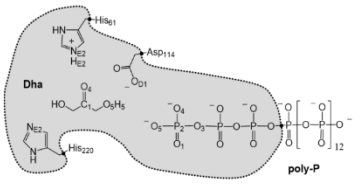


Figure 2. Cartoon representation of the kinase homodimer obtained by supeposition of PDB code crystal structures 1UN8 and 1UN9. The DhaK-domain is colored in pink and red, and the DhaL-domain is colored in purple and grey. Dha, phosphoaminophosphonic acid-adelynate ester (ANP), magnesium ions and phospholipid are shown in a ball and sticks representation.

Thereby, recalculation of the pKa values of the titratable residues was determined using the PROPKA program of Jensen *et al.*[49,50] According to the results, all residues were found at their standard protonation state in aqueous solution, except His61 that was double protonated. Afterwards, a 16 monomers poly-P was docked into the active site by placing a phosphate group in a position equivalent to the ATP for the chemical reaction to occur. Subsequently, the poly-P was slightly displaced by means of geometry optimizations and short Langevin-Verlet[51] MD (NVT) in order to avoid overlapping between the poly-P and the protein and to get an adequate protein-poly-P complex for the chemical steps to take place. This calculations were carried out at 300 K of temperature using the fDYNAMO library.[52] The resulting structure showed a better conformation pose of the poly-P establishing reasonable interactions with the residues of the protein and Mg cations. The poly-P was further equilibrated by means of 10 ns of MD simulations fixing all the protein atoms and constraining the distance between Dha and the poly-P. The mutated enzyme was generated from this structure by replacing Glu526 with a lysine residue. Then, 29 and 27 counter ions Na^+ were placed into optimal electrostatic positions around the wild type and the mutant, respectively. Afterwards, the systems were solvated with pre-equilibrated orthorhombic boxes of water molecules with dimensions 100 x 80 x 80 $Å^3$ centred on the centre of mass of poly-P. Water molecules with an oxygen atom lying within 2.8 Å of any heavy atom of poly-P, Dha, phospholipid, ions or protein were removed. Later, optimizations using the conjugated gradient algorithm were performed within the NAMD[53] parallel molecular dynamics code using the CHARMM[54-56] force field while water molecules were described with TIP3P force field[57]. Cutoffs for the nonbonding interactions were applied using a switching function, within a radius range from 14.0 to 16.0 Å, employing in all the simulations periodic boundary conditions.

The resulting systems were taken as initial structures for the study of the chemical reaction by means of QM/MM calculations. The Dha molecule, three phosphate groups from poly-P and the side chain of residues involved in the reaction mechanism (His-61, Asp-114 and His-220) were described quantum mechanically while the rest of the atoms of the system and water molecules were treated by means of OPLS-AA[58,59] and TIP3P[57] force fields, respectively. The same treatment as in the previous MD simulations was applied for the non-bonding interactions. To saturate the valence of the QM/MM frontier atoms, the link

atom procedure was employed.[60] Thus, quantum link atoms where placed between the third phosphate bridge oxygen and the fourth phosphorous atom in the case of poly-P and between C_α and C_β, in Asp-114 and both histidine residues (see Scheme 2). Therefore, the QM part involves 54 atoms with a total charge of -4. For all simulations, atoms belonging to molecules found at a distance less or equal than 25 Å from the poly-P were defined as flexible. The rest of the atoms were kept frozen. Before exploring the corresponding PES, series of MM and QM/MM L-BFGS-B optimizations were applied to fully relax the systems in a reactant-like conformation.



Scheme 2. Schematic representation of the active site of the DHAK. Grey region contains atoms treated quantum mechanically. Link atoms are represented as dots.

The PM3 semiempirical method,[61] implemented in fDYNAMO library, was employed to describe the QM region during the QM/MM simulations. PM3 has been proved to produce a considerable stabilization in the energy of phosphorane species owing to the use of a minimal valence basis and being normally 3-4 orders of magnitude faster than DFT methods.[62] Consequently, it has been employed widely to model phosphorus and phosphate groups successfully.[63-66] In fact, the semiempirical Hamiltonian PM3 was already used in our group to describe the QM sub-set of atoms in our previous QM/MM studies of the phosphate transfer reaction between ATP and Dha in aqueous solution and catalysed by DHAK.[4,23] In addition, the B3LYP functional, with the 6-31G(d,p) basis set, was also used to treat the QM region of the system combining fDYNAMO library with Gaussian09 program.[67] B3LYP functional has been successfully employed by Fernandes and co-workers to study the reaction mechanism of mycobacterium tuberculosis glutamine synthetase (mtGS), whose first step involves the phosphate group transfer from ATP to a glutamine residue.[68] The *substrate-assisted* and the *asp-assisted mechanisms* for the phosphate transfer from poly-P to Dha in wild type and the E526K mutant of *C. freundii* DHAK was explored by generating the PESs corresponding to each chemical step scanning the appropriate combination of the interatomic distances (see next section for details). Stationary point structures (reactants, intermediates, TSs and products) were refined and characterized, guided by a micro-macro iterations scheme.[69]

Results and discussion

PM3/MM results. Figure 3 shows the active site of the wild type and the E526K mutant in a reactant state conformation. As explained in previous section, these structures were obtained after series of MD simulations and QM/MM optimizations in both systems. It can be noted how the poly-P is well posed for the phosphate transfer to Dha, in both systems. A detailed analysis of the structures suggests that the binding of the poly-P to the protein in the vicinities of the active site is slightly better arranged in the mutant than in the wild type, in agreement with our previous study of the binding step.[31]

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

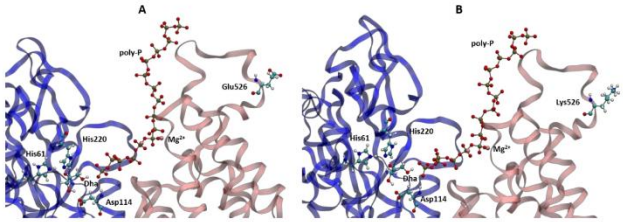


Figure 3. Active site of the wild-type (A) and the E526K mutant of *C. freundii* DHAK (B) in the reactants state conformation. Key residues (Asp114, His220, His61 and Glu/Lys526), Dha, poly-P, and Mg^{2+} ions are displayed in ball and stick representation.

In the initial catalytic step of any of the two proposed mechanisms,[21,22] the Dha is anchored to the enzyme through the formation of a covalent bond between the nitrogen atom NE2 of the His220 and the carbon C1 of the Dha. Then, the double protonated His61 transfers the HE2 proton to the oxygen atom O4 of Dha obtaining the required intermediate, I2. The PES of this step has been generated using as distinguished reaction coordinates the distance between the nitrogen atom of His220 and the carbon atom of Dha, $d(NE2_{His220} - C1)$, and the antisymmetric combination of the bond-breaking and bond-forming distances that describes the proton transfer from His61 to Dha, $d(NE2_{His61} - HE2_{His61}) - d(O_4 - HE2_{His61})$. The corresponding PESs obtained for wild type and mutant are displayed in Figure 4 and a schematic representation of the reaction mechanism derived from these surfaces are shown in Scheme 3.

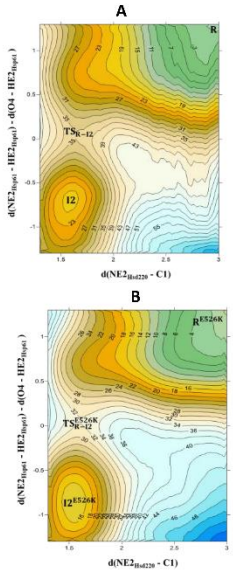
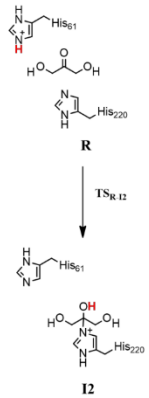


Figure 4. PM3/MM PESs of the initial catalytic step of the phosphorylation reaction mechanism from poly-P to Dha in *C. freundii* DHAK. Results are shown for the wild-type (A) and the E526K mutant (B). Distances on axis are in Å, and values of isoenergetic lines in kcal·mol^{-1}.



Scheme 3. Schematic representation of the concerted mechanism for the covalent bond formation between Dha and the enzyme; the wild type and the E526K mutant of the *C. freundii* DHAK.

Figure 4 shows that the transformation from R to I2 takes place through a concerted mechanism in both systems within only one TS, $TS_{R\text{-}I2}$. Interestingly, the energy barrier is slightly lower in the E526K mutant than in the wild type (32 kcal·mol^{-1} and 35 kcal·mol^{-1} respectively), which would be in agreement with its experimentally observed catalytic activity, not detected in the wild type.[31]

Thus, once this intermediate I2 is reached, two possible phosphorylation reaction paths can take place, the *asp-assisted mechanism*, and the *substrate-assisted mechanism*. In the former, the catalytic Asp114 acts as a basis accepting the proton transferred from Dha and activating it for the nucleophilic attack to the poly-P. This mechanism would be equivalent to that experimentally proposed by Shi and co-workers in DHAK from *E. coli*.[22] In the *substrate-assisted mechanism* the poly-P molecule directly abstracts the proton from Dha. De Vivo et. al supported this mechanism in the theoretical study of the phosphate transfer reaction from ATP to a serine residue in CDK2,[21] as well as our previous study of the phosphate transfer reaction from ATP to Dha on wild type DHAK.[23] Both reaction mechanisms are explored for the phosphate transfer reaction from poly-P, instead of ATP, to Dha on wild type and E526K mutant in the present study.

The *asp-assisted mechanism*. The PESs to explore this mechanism were generated within two antisymmetric combinations of inter-atomic distances describing the proton transfer from Dha to Asp114, $d(O5_{Dha} - H5_{Dha}) - d(H5_{Dha} - OD1_{Asp114})$, and the phosphate transfer from the poly-P to Dha, $d(O3_{poly\text{-}P} - P2_{poly\text{-}P}) - d(P2_{poly\text{-}P} - O5_{Dha})$. Figure 5 shows the resulting PESs for the wild type and the E526K mutant while a schematic representation of the resulting mechanism is presented in the Scheme 4.

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

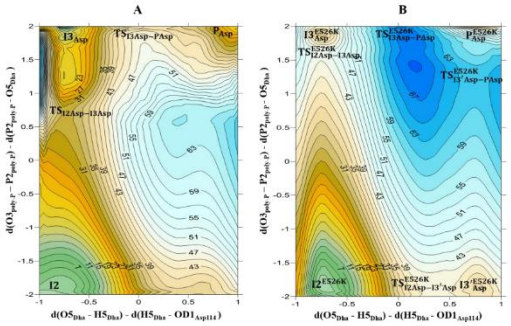
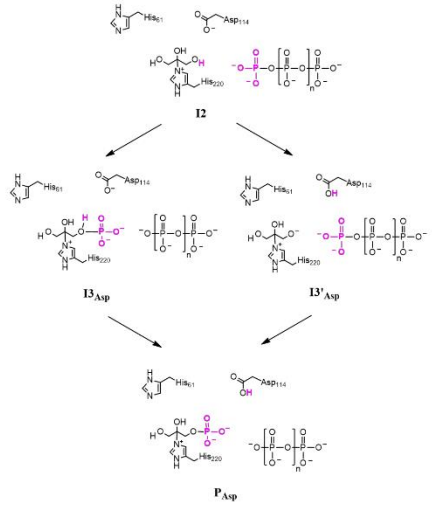


Figure 5. PM3/MM PESs of the catalytic process of the *aspartate-assisted* phosphorylation reaction mechanism from poly-P to Dha in *C. freundii* DHAK. Results are shown for the wild-type (A) and the E526K mutant (B). Distances on axis are in Å, and values of isoenergetic lines are in kcal·mol^{-1}.



Scheme 4. Schematic representation of the *asp-assisted mechanism* for the phosphorylation reaction from I2 to P_{Asp}.

According to the PES presented in Figure 5A, the phosphate would be transferred from poly-P to Dha to produce a stable intermediate in the wild-type enzyme, $I3_{Asp}$. Then, the proton is transferred from Dha to Asp thus forming the products P_{Asp}. This second step presents an energy barrier (39 kcal·mol^{-1}) significantly higher than the one corresponding to the first step (32 kcal·mol^{-1}). As shown in Figure 5B, the reaction in the mutant can take place through a similar mechanism where the phosphate transfer precedes the proton transfer through a stable intermediate, $I3_{Asp}^{E526K}$. But the reaction could also proceed through a metastable intermediate $I3'^{E526K}_{Asp}$. Nevertheless, despite the existence of this shallow minimum, this second path can take place in a very asynchronous but concerted mechanism controlled just by the $TS_{I3'Asp-PAsp}^{E526K}$. In any case, both alternative reaction pathways take place through TSs that appear at noticeable higher energies than those obtained in the wild type enzyme, 42 and 66 kcal·mol^{-1} for the reaction where the phosphate transfer precedes the proton transfer, and 62 kcal·mol^{-1} in the alternative asynchronous concerted reaction path, by comparison with the 39 kcal·mol^{-1} of the rate limiting TS in the wild-type. The analysis of the reaction coordinates of the TS quadratic regions indicate quite unusual values for the TS, especially for the phosphate transfer: +0.6 Å and +1.5 Å for the $TS_{I2Asp-I3Asp}$ and the $TS_{I2Asp-I3Asp}^{E526K}$, respectively. Moreover, it must be noticed that the located intermediates, $I3_{Asp}$ and $I3_{Asp}^{E526K}$, present a significantly large interatomic distances between the phosphate and its acceptor atom (1.92 and 2.11 Å in the wild type and mutant, respectively). This must be related with the fact that the transferring proton, $H5_{Dha}$, is still bounded to the oxygen donor atom, $O5_{Dha}$, in both systems. A detailed analysis of the interatomic distances of the stationary points listed in Table 1 confirms the comments on the TSs and the character of the $I3_{Asp}$ and $I3_{Asp}^{E526K}$ intermediates.

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

Table 1. Key interatomic distances (in Å) of the stationary points located along the *asp-assisted mechanism* of the phosphate transfer between poly-P and Dha obtained at PM3/MM level in the A) wild type, and B) E526K mutant.

A) Wild type

Atomic distances	I2	$I2_{I2Asp-I3Asp}$	$I3_{Asp}$	$I2_{I3Asp-PAsp}$	P_{Asp}
$O5_{Dha}$-$H5_{Dha}$	0.99	1.00	0.98	1.15	1.93
$H5_{Dha}$-$OD1_{Asp114}$	1.85	1.68	1.70	1.20	0.96
$O3_{poly-P}$-$P2_{poly-P}$	1.91	3.17	3.80	3.70	3.80
$P2_{poly-P}$-$O5_{Dha}$	3.92	2.56	1.92	1.90	1.89

B) E526K mutant

Atomic distances	$I2^{E526K}$	$TS^{E526K}_{I2Asp-I3Asp}$	$I3^{E526K}_{Asp}$	$TS^{E526K}_{I3Asp-PAsp}$	$TS^{E526K}_{I2Asp-I3'Asp}$	$I3'^{E526K}_{Asp}$	$TS^{E526K}_{I3'Asp-PAsp}$	P^{E526K}_{Asp}
$O5_{Dha}$-$H5_{Dha}$	0.99	0.99	1.00	1.19	1.50	1.73	1.77	1.97
$H5_{Dha}$-$OD1_{Asp114}$	1.77	1.76	1.73	1.29	1.06	0.97	0.99	1.03
$O3_{poly-P}$-$P2_{poly-P}$	1.95	3.81	3.98	3.98	2.01	2.03	3.49	3.88
$P2_{poly-P}$-$O5_{Dha}$	3.85	2.25	2.11	2.11	3.82	3.84	2.51	1.99

The *substrate-assisted mechanism*. The PESs to explore this mechanism were generated with the antisymmetric combinations of distances describing the proton and the phosphate transfer, $d(O5_{Dha} - H5_{Dha}) - d(H5_{Dha} - O5_{poly-P})$ and $d(O3_{poly-P} - P2_{poly-P}) - d(P2_{poly-P} - O5_{Dha})$, respectively. The resulting PESs are displayed in Figure 6 and the schematic representation of the mechanism is presented in Scheme 6. In this case, unlike in the case of the exploration of the *asp-assisted mechanism*, similar reaction paths were obtained for the wild-type and the mutant, as revealed by Figure 6A and 6B, and the interatomic distances listed in Table 2. Analysis of the surfaces suggests that in both systems the proton abstraction of Dha by the poly-P involving the formation of the intermediate I3 occurs first, and it is followed by the phosphate transfer. In this case the quadratic region of all the TSs located on Figure 6A and 6B describe quite synchronous breaking and forming bonds (the reaction coordinate are around 0 Å in all the 4 located TSs). This is confirmed by analysis of the interatomic distances of the stationary points listed in Table 2.

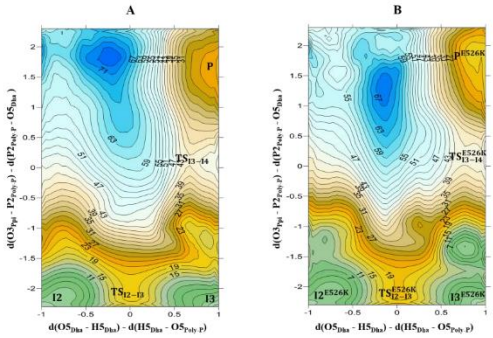
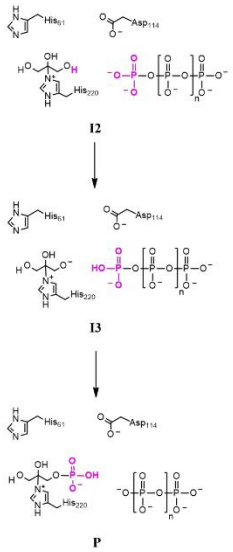


Figure 6. PM3/MM PESs of the catalytic process of the *substrate-assisted* phosphorylation reaction mechanism from poly-P to Dha in *C. freundii* DHAK. Results are shown for the wild-type (A) and the E526K mutant (B). Distances on axis are in Å, and values of isoenergetic lines are in kcal·mol^{-1}.



Scheme 6. Schematic representation of the *substrate-assisted mechanism* for the phosphorylation reaction from I2 to P for wild and mutant enzyme.

Table 2. Key interatomic distances (in Å) of the stationary points located along the *substrate-assisted mechanism* of the phosphate transfer between poly-P and Dha obtained at PM3/MM level in the A) wild type, and B) E526K mutant.

A) Wild type

Atomic distances	I2	TS_{I2-I3}	I3	TS_{I3-P}	P
$O5_{Dha}$-$H5_{Dha}$	1.00	1.24	1.89	1.70	1.95
$H5_{Dha}$-$O5_{poly-P}$	1.77	1.26	0.93	0.96	0.95
$O3_{poly-P}$-$P2_{poly-P}$	1.85	1.89	1.76	2.56	3.33
$P2_{poly-P}$-$O5_{Dha}$	4.00	3.81	4.09	2.49	1.75

B) E526K mutant

Atomic distances	$I2^{E526K}$	TS_{I2-I3}^{E526K}	$I3^{E526K}$	TS_{I3-P}^{E526K}	P^{E526K}
$O5_{Dha}$-$H5_{Dha}$	0.98	1.19	1.69	1.73	1.85
$H5_{Dha}$-$O5_{poly-P}$	1.71	1.13	0.98	0.96	0.95
$O3_{poly-P}$-$P2_{poly-P}$	1.79	1.76	1.73	2.59	3.75
$P2_{poly-P}$-$O5_{Dha}$	3.81	3.50	3.85	2.52	1.77

Interestingly, the energy barriers are 20 and 42 kcal·mol^{-1} for the reaction in the wild-type, and 20 and 40 kcal·mol^{-1} for the reaction in the mutant. These values are significantly lower than those reported for the *asp-assisted mechanism* in the wild type and in the mutated enzyme. This trend is in agreement with our previous study of the phosphate transfer from ATP to Dha catalyzed by DHAK that also showed the *substrate-assisted mechanism* as the most favorable one.[23]

Keeping in mind the negative charge of the transferring phosphate, it is not surprising that the activation of nucleophile by an aspartate residue, increasing the negative charge on the $O5_{Dha}$ atom, does not favor the phosphate transfer. On the contrary, a proton transfer from Dha to the poly-P can favor the reaction from an electrostatic point of view. Therefore, the much more expensive B3LYP/MM calculations have been focused just on the *substrate-assisted mechanism*.

DFT/MM Results. The stationary points corresponding to the phosphate transfer from poly-P to the Dha covalently bound to the enzyme (I2 intermediate), through the *substrate-assisted mechanism,* was located and characterized at B3LYP/MM level from initial guess structures of the wild-type and the mutant obtained from the PM3/MM PESs. Key interatomic distances of the located stationary points, together with the relative energies to I2, are reported in Table 3. The first relevant result is that the reaction, either in the wild-type or the E526K mutant, takes place in a single step with high but similar energy barriers: 66.3 and 71.2 kcal·mol^{-1}, respectively.

Analysis of the interatomic distances of the corresponding TSs, TS_{I2-P} and TS_{I2-P}^{E526K} schematically presented in Figure 7, shows values of the antisymmetric combination of the breaking and forming bonds that are +0.58 Å and +0.53 Å for the transfer of the proton and phosphate in wild-type, and +0.64 Å and +0.52 Å in the mutant, respectively. Thus, the results at B3LYP/MM level describe the two transfers in an advanced stage of the process in both TSs. In

both reactions, the proton is almost completely transferred in the TS and, according to the interatomic distances between the phosphorous atom and the donor and acceptor oxygen atoms (2.75 and 2.22 Å in the wild type, and 2.69 and 2.17 Å in the mutant), they can be considered as quite dissociative. It appears that the interactions with a Na ion, water molecules and with the residues Gly113 of the active site stabilize the transferring phosphate in both TSs. The O···P···O angle in both systems is also quite similar at the TSs (153 and 152 degrees in the wild type and in the mutant, respectively). Charge analysis on key atoms of the TSs show how charge on oxygen donor atom is only slightly higher in the mutant (-1.049 a.u.) than in the wild-type (-1.044 a.u). The oxygen acceptor atom is less negatively charged in the mutant (-0.615 a.u.) than in the wild-type (-0.662 a.u.). In all, the sum of the charges of the protonated transferring phosphate group (adding the charge of the transferring proton since it is almost completely transferred to the $O5_{poly\text{-}P}$ atom in both TSs) is -0.555 and -0.568 a.u. in the wild-type and in the mutant, respectively. Tables containing the full list of atomic charges on the two TSs are deposited in the Supporting Information.

Analysis of the structures can be also carried out by inspection of Figure 8, where the superposition of the I2 and TS structures in the wild-type and in the mutant are presented. As observed, the structure of the proteins in both states are almost coincident, revealing insignificant effects of the E526K mutation on the chemical step of the full process.

Table 3. Key interatomic distances (in Å) and relative potential energies (E, in kcal·mol^{-1}), for the stationary points localized at B3LYP/MM level for the phosphoryl transfer step through the *substrate-assisted mechanism* in A) the wild type enzyme, and (B) the E526K mutant.

A) Wild type

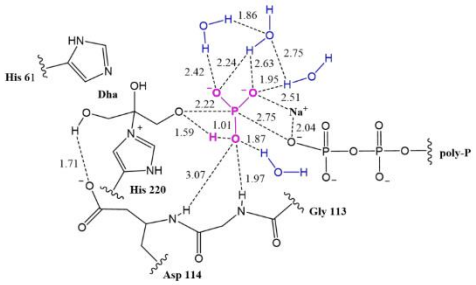
	$(O5_{Dha}\text{-}H5_{Dha})$	$(H5_{Dha}\text{-}O5_{poly\text{-}P})$	$(P2_{poly\text{-}P}\text{-}O3_{poly\text{-}P})$	$(P2_{poly\text{-}P}\text{-}O5_{Dha})$	E
I2	1.04	1.62	1.71	3.77	0
$\mathbf{TS_{I2-P}}$	1.59	1.01	2.75	2.22	66.3
P	2.01	0.97	3.42	1.74	57.8

B) E526K mutant

$\mathbf{I2^{E526K}}$	1.07	1.52	1.71	3.75	0
$\mathbf{TS_{I2-P}^{E526K}}$	1.64	1.00	2.69	2.17	71.17
$\mathbf{P^{E526K}}$	1.99	0.97	3.27	1.75	64.97

International Journal of QUANTUM CHEMISTRY WWW.Q-CHEM.ORG FULL PAPER

(A)



(B)

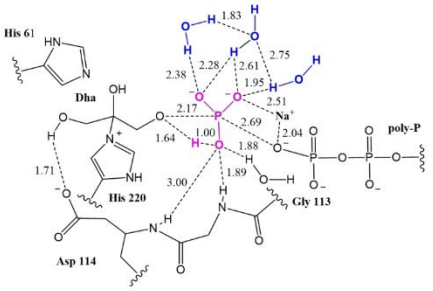


Figure 7. B3LYP/MM optimized structure corresponding to the transition state of the phosphoryl transfer step for the *substrate-assisted mechanism* obtained (A) in the wild type enzyme, and (B) in the E526K mutant.

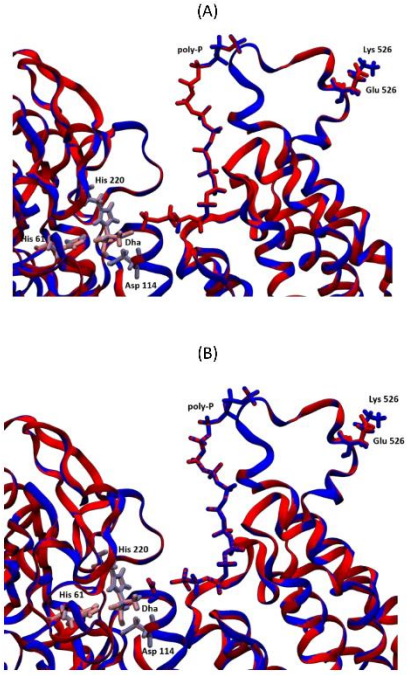


Figure 8. Superposition of B3LYP/MM optimized structures in the wild-type (in red and pink) and in the E526K mutant (in blue and grey) corresponding to I2 intermediate (A) and the transition state (B) of the phosphoryl transfer step for the *substrate-assisted mechanism*.

Conclusions

In this paper we present a computational study of the molecular mechanism of the phosphoryl transfer reaction from an inorganic polyphosphate to Dha catalyzed by wild type DHAK from *C. freundii*, and by a mutant, Glu526Lys, that has shown activity with poly-P as phosphoryl donor.[31] The study is first based on MD simulations to equilibrate both systems. Then, PESs corresponding to every single step of the process were explored for two different reaction mechanisms, an *asp-assisted mechanism* and a *substrate-assisted mechanism*, using hybrid QM/MM potentials, with the QM region described at the PM3 semiempirical level. The comparison of the obtained PESs for the central catalytic step of the reaction (i.e. the phosphate transfer from the inorganic poly-P to the Dha molecule from the structure where Dha is covalently bound to His220 residue of the protein, namely I2) shows that the *substrate-assisted mechanism* is kinetically more favorable than the asp-*assisted mechanism*, in both the wild-type and the mutant. It appears that the transfer of the

proton from Dha to Asp114 does not really activate its oxygen acceptor atom since it becomes more negative. In contrast, the transfer of the proton to the phosphate group reduces its negative charge and, consequently, facilitates the approach to the acceptor oxygen atom of the Dha. Then, the key stationary points of the *substrate-assisted mechanism* were localized with the B3LYP functional to describe the QM sub-set of atoms. The DFT/MM calculations clearly indicate that the phosphate transfer can take place in a single step, with quite dissociative TSs. Nevertheless, the energy barriers appear to be quite high. A note of caution must be introduced at this point since, due to the fact that the study is based on exploration of PESs, the detected differences between located stationary point structures could be due just to the different conformations on the located structures, probably not significant taking into account the large systems under study. Then, these results of potential energies of single molecules optimizations have to be considered with caution. A proper sampling of conformations on the involved states, i.e. the calculation of free energy barriers from statistical simulations, can render significantly different results since, as observed in our pioneering QM/MM studies of enzyme catalyzed reactions and supported by single molecule kinetic experiments,[70] the differences between the rate constant of single molecule[71,72] can oscillate in several orders of magnitude around. A much more computationally demanding sampling of the conformational space through the use of hybrid DFT/MM molecular dynamics could perfectly render equivalent average structures for the wild type and the mutant.

Anyway, analysis of the TS structures located at DFT/MM level suggest that the *substrate-assisted mechanism*, that a priori could not be considered as a catalytic process since it could also take place in a reference reaction in solution, is favored by the interactions with the residues of the active site that stabilize the negatively charged phosphate group in the TS. Superposition of the reactants (the state where the Dha is anchored to the protein through His220) and TS structures in the wild-type and in the mutant shows that the structure of the proteins in both states are almost coincident, suggesting insignificant effects of the E526K mutation on the chemical step of the full process. Thus, this geometrical analysis together with the similar energy barriers obtained in both systems, confirm that this mutation does not have an effect on the chemical reaction step of the process.

The present results, that are part of a project devoted to modify the phosphoryl donor specificity of this enzyme from ATP to a poly-P, complement and confirm the conclusions obtained in our previous study of the binding of poly-P to wild-type and E526K DHAK from *C.freundii*.[31] The measured activity of some mutants as the one tested in this computational study where the replacement of residues are done on positions far from the active site, can be associated to an improvement of the formation of the protein-substrate complex. Further studies could be done in order to check whether this long distance mutation can be combined with other mutations on the proximity of the active site. In this regard, the present study of the active centre during the chemical reaction will be of great interest to select the appropriate residues to be replaced.

Acknowledgements

This work was supported by the Spanish Ministerio de Economía y Competitividad for project CTQ2015-66223-C2, Universitat Jaume I (project P1•1B2014-26), Generalitat Valenciana (PROMETEOII/2014/022) Universitat Jaume I (project P1•1B2014-26 and P1•1B2013-58). V.M. is grateful to the University of Bath for the award of a David Parkin Visiting Professorship. Authors acknowledge computational resources from the Servei d'Informàtica of Universitat Jaume I.

Supporting information

Cartesian coordinates of the QM atoms of the TSs localized at B3LYP/MM level, list of atomic charges on key atoms of TSs located at B3LYP/MM level.

References and Notes

(1) Hengge, A. C. In *Encyclopedia of Life Sciences*; John Wiley & Sons, London, 2004.
(2) Kamerlin, S. C. L.; Williams, N. H.; Warshel, A. *J. Org. Chem.* **2008**, *73*, 6960-6969.
(3) Kamerlin, S. C. L. *J. Org. Chem.* **2011**, *76*, 9228-9238.
(4) Bordes, I.; Ruiz-Pernía, J. J.; Castillo, R.; Moliner, V. *Org. Biomol. Chem.* **2015**, *13*, 10179-10190.
(5) Kamerlin, S. C. L.; Wilkie, J. *Org. Biomol. Chem.* **2007**, *5*, 2098-2108.
(6) Kamerlin, S. C. L.; Florian, J.; Warshel, A. *ChemPhysChem* **2008**, *9*, 1767-1773.
(7) Rosta, E.; Kamerlin, S. C. L.; Warshel, A. *Biochemistry* **2008**, *47*, 3725-3735.
(8) Ptacek, J.; Snyder, M. *Trends Genet.* **2006**, *22*, 545-554.
(9) Kinnings, S. L.; Jackson, R. M. *J. Chem. Inf. Model.* **2009**, *49*, 318-329.
(10) Cheek, S.; Ginalski, K.; Zhang, H.; Grishin, N. V. *BMC Struct. Biol.* **2005**, *5*, 1-19.
(11) Deshmukh, K.; Anamika, K.; Srinivasan, N. *Prog. Biophys. Mol. Biol.* **2010**, *102*, 1-15.
(12) Mijakovic, I.; Macek, B. *Fems Microbiol. Rev.* **2012**, *36*, 877-892.
(13) Rakshambikai, R.; Manoharan, M.; Gnanavel, M.; Srinivasan, N. *Rsc Adv.* **2015**, *5*, 25132-25148.
(14) Garcia-Garcia, T.; Poncet, S.; Derouiche, A.; Shi, L.; Mijakovic, I.; Noirot-Gros, M.-F. *Front. Microbiol.* **2016**, *7*.
(15) Pérez-Gallegos, A.; García-Viloca, M.; González-Lafont, A.; Lluch, J. M. *Phys. Chem. Chem. Phys.* **2015**, *17*, 3497-3511.
(16) Diaz, N.; Field, M. J. *J. Am. Chem. Soc.* **2004**, *126*, 529-542.
(17) Cheng, Y. H.; Zhang, Y. K.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1553-1562.
(18) Pérez-Gallegos, A.; García-Viloca, M.; González-Lafont, A.; Lluch, J. M. *ACS Catal.* **2015**, *5*, 4897-4912.
(19) Ojeda-May, P.; Li, Y.; Ovchinnikov, V.; Nam, K. *J. Am. Chem. Soc.* **2015**, *137*, 12454-12457.
(20) Hart, J. C.; Sheppard, D. W.; Hillier, I. H.; Burton, N. A. *Chem. Commun.* **1999**, 79-80.
(21) De Vivo, M.; Cavalli, A.; Carloni, P.; Recanatini, M. *Chem. - Eur. J.* **2007**, *13*, 8437-8444.
(22) Shi, R.; McDonald, L.; Cui, Q.; Matte, A.; Cygler, M.; Ekiel, I. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1302-1307.
(23) Bordes, I.; Castillo, R.; Moliner, V. *ASAP J. Phys. Chem. B* **2017**.
(24) Zurbriggen, A.; Jeckelmann, J.-M.; Christen, S.; Bieniossek, C.; Baumann, U.; Erni, B. *J. Biol. Chem.* **2008**, *283*, 35789-35796.
(25) García-Alles, L. F.; Siebolo, C.; Nyffeler, T. L.; Flukiger-Bruhwiler, K.; Schneider, P.; Burgi, H. B.; Baumann, U.; Erni, B. *Biochemistry* **2004**, *43*, 13037-13045.
(26) Bachler, C.; Flukiger-Bruhwiler, K.; Schneider, P.; Bahler, P.; Erni, B. *J. Biol. Chem.* **2005**, *280*, 18321-18325.
(27) Erni, B.; Siebold, C.; Christen, S.; Srinivas, A.; Oberholzer, A.; Baumann, U. *Cell. Mol. Life Sci.* **2006**, *63*, 890-900.
(28) Shi, R.; McDonald, L.; Cygler, M.; Ekiel, I. *Struct.* **2014**, *22*, 478-487.
(29) Raynaud, C.; Lee, J.; Sarcabal, P.; Croux, C.; Meynial-Salles, I.; Soucaille, P. *J.f Bacteriol.* **2011**, *193*, 3127-3134.
(30) Enders, D.; Voith, M.; Lenzen, A. *Angew. Chem., Int. Ed.* **2005**, *44*, 1304-1325.
(31) Sánchez-Moreno, I.; Bordes, I.; Castillo, R.; Ruiz-Pernía, J. J.; Moliner, V.; García-Junceda, E. *Int. J. Mol. Sci.* **2015**, *16*, 26073.
(32) Sánchez-Moreno, I.; Iturrate, L.; Doyagueez, E. G.; Antonio Martínez, J.; Fernández-Mayoralas, A.; García-Junceda, E. *Adv. Synth. Catal.* **2009**, *351*, 2967-2975.
(33) Fesko, K.; Gruber-Khadjawi, M. *ChemCatChem* **2013**, *5*, 1248-1272.
(34) Iturrate, L.; Sánchez-Moreno, I.; Oroz-Guinea, I.; Pérez-Gil, J.; García-Junceda, E. *Chem. - Eur. J.* **2010**, *16*, 4018-4030.

(35) Schumperli, M.; Pellaux, R.; Panke, S. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 33-45.
(36) Buchholz, K; Volker, K.; Bornscheuer, U. T. *Biocatalysts and Enzyme Technology*, 2nd Edition, Wiley-VCH Verlag & Co., Weinheim, Germany, 2012.
(37) Windle, C. L.; Mueller, M.; Nelson, A.; Berry, A. *Curr. Opin. Chem. Biol.* **2014**, *19*, 25-33.
(38) Sánchez-Moreno, I.; García-García, J. F.; Bastida, A.; García-Junceda, E. *Chem. Commun.* **2004**, *14*, 1634-1635.
(39) Iwamoto, S.; Motomura, K.; Shinoda, Y.; Urata, M.; Kato, J.; Takiguchi, N.; Ohtake, H.; Hirota, R.; Kuroda, A. *Appl. Environ. Microbiol.* **2007**, *73*, 5676-5678.
(40) Restiawaty, E.; Iwasa, Y.; Maya, S.; Honda, K.; Omasa, T.; Hirota, R.; Kuroda, A.; Ohtake, H. *Process Biochem.* **2011**, *46*, 1747-1752.
(41) Kameda, A.; Shiba, T.; Kawazoe, Y.; Satoh, Y.; Ihara, Y.; Munekata, M.; Ishige, K.; Noguchi, T. *J. Biosci. Bioeng.* **2001**, *91*, 557-563.
(42) Sato, M.; Masuda, Y.; Kirimura, K.; Kino, K. *J. Biosci. Bioeng.* **2007**, *103*, 179-184.
(43) Brown, M. R. W.; Kornberg, A. *Trends Biochem Sci* **2008**, *33*, 284-290.
(44) Rao, N. N.; Gomez-Garcia, M. R.; Kornberg, A. *Annu. Rev. Biochem.* **2009**, *78*, 605-647.
(45) Kornberg, A.; Rao, N. N.; Ault-Riche, D. *Annu. Rev. Biochem.* **1999**, *68*, 89-125.
(46) Siebold, C.; Arnold, I.; García-Alles, L. F.; Baumann, U.; Erni, B. *J. Biol. Chem.* **2003**, *278*, 48236-48244.
(47) Oberholzer, A. E.; Schneider, P.; Baumann, U.; Erni, B. *J. Mol. Biol.* **2006**, *359*, 539-545.
(48) Schaftenaar, G.; Noordik, J. H. . *Comput. - Aided Mol. Des.* **2000**, *14*, 123-134.
(49) Bertran Rusca, J.; Branchadell Gallo, V.; Moreno Ferrer, M.; Sodupe Roure, M. *Química cuántica. Fundamentos y aplicaciones computacionales*, 2nd Edition, Madrid, 2002.
(50) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704-721.
(51) Verlet, L. *Phys. Rev.* **1967**, *159*, 98-+.
(52) Field, M. *A Practical Introduction to the Simulation of Molecular Systems*, 2nd Edition, Cambridge University Press, Cambridge, 2007.
(53) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781-1802.
(54) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671-690.
(55) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
(56) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400-1415.
(57) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.
(58) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666.
(59) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810-2819.
(60) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700-733.
(61) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209-220.
(62) Nam, K.; Cui, Q.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 486-504.
(63) Corrie, J. E. T.; Barth, A.; Munasinghe, V. R. N.; Trentham, D. R.; Hutter, M. C. *J. Am. Chem. Soc.* **2003**, *125*, 8546-8554.
(64) Brandt, W.; Dessoy, M. A.; Fulhorst, M.; Gao, W. Y.; Zenk, M. H.; Wessjohann, L. A. *ChemBioChem* **2004**, *5*, 311-323.
(65) Hand, C. E.; Honek, J. F. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 183-188.
(66) Plotnikov, N. V.; Prasad, B. R.; Chakrabarty, S.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **2013**, *117*, 12807-12819.
(67) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; Nakatsuji, H.; Cariaco, M.; Li, X.; Hratchian, H.;

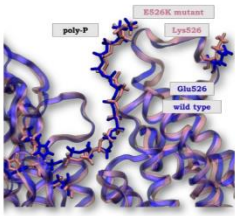
Izmaylov, A.; Bloino, J.; Zheng, G.; Sonnenberg, J.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr, J.; Peralta, J.; Ogliaro, F.; Bearpark, M.; Heyd, J.; Brothers, E.; Kudin, K.; Staroverov, V.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.; Iyengar, S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.; Klene, M.; Knox, J.; Cross, J.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; RL, M.; Morokuma, K.; Zakrzewski, V.; Voth, G.; Salvador, P.; Dannenberg, J.; Dapprich, S.; Daniels, A.; Farkas, O.; Foresman, J.; Ortiz , J.; Cioslowski, J. a.; Fox, D. *Gaussian, Inc., Wallingford CT* **2009**.
(68) Moreira, C.; Ramos, M. J.; Fernandes, P. A. *Chem. - Eur. J.* **2016**, *22*, 9218-9225.
(69) Marti, S.; Moliner, V.; Tuñón, I. *J. Chem. Theory Comput.* **2005**, *1*, 1008-1016.
(70) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323-1331.
(71) Xue, Q. F.; Yeung, E. S. *Nature* **1995**, *373*, 681-683.
(72) Tan, W. H.; Yeung, E. S. *Anal. Chem.* **1997**, *69*, 4242-4248.

GRAPHICAL ABSTRACT

Isabel Bordes Pastor, Eduardo García-Junceda, Raquel Castillo Solsona, Vicent Moliner Ibáñez

Computational Study of the Phosphoryl Donor Activity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphosphate

Herein it is presented a QM/MM theoretical study of the molecular mechanism for the phosphoryl transfer reaction from an inorganic polyphosphate to Dha catalyzed by DHAK from *C.freundii* (wild-type and an active experimentally measured mutant), as part of a project devoted to modify the phosphoryl donor specificity of this enzyme. The similar energy barriers obtained in both systems confirm our previous studies that this mutation improve the binding step of the process.



Supporting Information

Computational Study of the Phosphoryl Donor Activity of Dihydroxyacetone Kinase from ATP to Inorganic Polyphoshate

Isabel Bordes Pastor,[1] Eduardo García-Junceda,[2] Raquel Castillo Solsona,[1,*] Vicent Moliner Ibáñez[1,*]

Correspondence to:

Raquel Castillo Solsona (E-mail: rcastill@uji.es) tel: +34964728097

Vicent Moliner Ibáñez (E-mail: moliner@uji.es) tel: +34964728084

[1] *Isabel Bordes Pastor, Raquel Castillo Solsona, Vicent Moliner Ibáñez*
Departament de Química Física i Analítica, Universitat Jaume I, Castelló, Spain, 12071
[2] *Eduardo García-Junceda*
Departamento de Química Bioorgánica, Instituto de Química Orgánica General, CSIC, Juan de la Cierva 3, Madrid, Spain, 28006

Table S1. Cartesian coordinates of the QM atoms of the TS localized at B3LYP/MM level of the *substrate-assisted mechanism* in the wild type enzyme.

DHA			
$C1_{Dha}$	-0.20448	-2.72954	4.37691
$C2_{Dha}$	1.29355	-2.41636	4.71958
$H21_{Dha}$	1.73460	-3.40182	4.94774
$H22_{Dha}$	1.29002	-1.83492	5.65412
$C3_{Dha}$	-0.37011	-3.61573	3.12463
$H31_{Dha}$	0.37247	-3.28667	2.38420
$H32_{Dha}$	-1.37146	-3.43140	2.69184
$O4_{Dha}$	-0.82532	-3.27713	5.48822
$O5_{Dha}$	1.94271	-1.76005	3.67296
$H5_{Dha}$	3.43642	-2.09664	3.24831
$O6_{Dha}$	-0.20929	-4.95286	3.51750
$H6_{Dha}$	0.03977	-5.50278	2.73454
POLY-P			
$O1_{poly\text{-}P}$	3.61869	0.12918	5.18274
$P2_{poly\text{-}P}$	3.51287	-0.18740	3.72877
$O3_{poly\text{-}P}$	6.11671	0.68159	3.68089
$O4_{poly\text{-}P}$	2.95403	0.73864	2.69810
$O5_{poly\text{-}P}$	4.25555	-1.50622	3.17463
$P6_{poly\text{-}P}$	7.65303	1.12838	3.57446
$O7_{poly\text{-}P}$	8.10670	1.12029	5.21934
$O8_{poly\text{-}P}$	8.48678	0.09530	2.82397
$O9_{poly\text{-}P}$	7.76466	2.58343	3.09799
$P10_{poly\text{-}P}$	9.40877	0.80886	6.34731
$O11_{poly\text{-}P}$	10.73144	1.56735	5.87856
$O12_{poly\text{-}P}$	8.76991	1.50971	7.57795
$O13_{poly\text{-}P}$	9.43241	-0.70593	6.49587
HIS 61			
CB_{His61}	-6.85799	-5.94647	4.92303
$HB1_{His61}$	-7.16370	-6.34866	3.94874
$HB2_{His61}$	-6.61650	-6.81175	5.54947
CG_{His61}	-5.61413	-5.12750	4.73339
$ND1_{His61}$	-5.30952	-4.54078	3.51604
$HD1_{His61}$	-5.94904	-4.47590	2.73808
$CE1_{His61}$	-4.09172	-3.93602	3.62251
$HE1_{His61}$	-3.62923	-3.38770	2.81462
$NE2_{His61}$	-3.57945	-4.11195	4.82482
$HE2_{His61}$	-1.69588	-3.67120	5.23294
$CD2_{His61}$	-4.52545	-4.84082	5.52268
$HD2_{His61}$	-4.38516	-5.10562	6.55862
HIS 220			
CB_{His220}	-2.26624	1.69292	5.56779
$HB1_{His220}$	-1.70597	1.69990	6.50840
$HB2_{His220}$	-1.94790	2.60150	5.06155
CG_{His220}	-1.76432	0.53893	4.75531
$ND1_{His220}$	-0.91820	0.73319	3.66932
$HD1_{His220}$	-0.56531	1.68672	3.28262

$CE1_{His220}$	-0.36410	-0.42687	3.31304
$HE1_{His220}$	0.42208	-0.55514	2.59194
$NE2_{His220}$	-0.84785	-1.38251	4.11648
$CD2_{His220}$	-1.73542	-0.79993	5.00792
$HD2_{His220}$	-2.21722	-1.38725	5.76635
ASP 114			
CB_{Asp114}	1.94051	-5.25047	0.34956
$HB1_{Asp114}$	2.00473	-4.67152	-0.57525
$HB2_{Asp114}$	1.74609	-4.56514	1.18219
CG_{Asp114}	0.72403	-6.19468	0.26521
$OD1_{Asp114}$	0.19182	-6.53922	1.38772
$OD2_{Asp114}$	0.25732	-6.54246	-0.84406

Table S2. Cartesian coordinates of the QM atoms of the TS localized at B3LYP/MM level of the *substrate-assisted mechanism* in the mutated enzyme.

DHA			
$C1_{Dha}$	-0.07479	-2.68209	4.35110
$C2_{Dha}$	1.43005	-2.36600	4.65771
$H21_{Dha}$	1.88634	-3.34884	4.87109
$H22_{Dha}$	1.45653	-1.78329	5.59081
$C3_{Dha}$	-0.27010	-3.54421	3.08430
$H31_{Dha}$	0.49316	-3.23387	2.35701
$H32_{Dha}$	-1.25929	-3.30664	2.64746
$O4_{Dha}$	-0.67308	-3.25984	5.46299
$O5_{Dha}$	2.037549	-1.70614	3.58874
$H5_{Dha}$	3.576659	-2.07227	3.15810
$O6_{Dha}$	-0.17640	-4.89796	3.45609
$H6_{Dha}$	0.13280	-5.44049	2.69264
POLY-P			
$O1_{poly\text{-}P}$	3.62199	0.15711	5.11157
$P2_{poly\text{-}P}$	3.56018	-0.15890	3.65448
$O3_{poly\text{-}P}$	6.11890	0.68195	3.62338
$O4_{poly\text{-}P}$	3.00340	0.75502	2.60928
$O5_{poly\text{-}P}$	4.35993	-1.45489	3.11082
$P6_{poly\text{-}P}$	7.66808	1.10794	3.53945
$O7_{poly\text{-}P}$	8.11086	1.03413	5.18477
$O8_{poly\text{-}P}$	8.48587	0.06899	2.78009
$O9_{poly\text{-}P}$	7.80111	2.56535	3.09366
$P10_{poly\text{-}P}$	9.41942	0.82030	6.32814
$O11_{poly\text{-}P}$	10.72814	1.54975	5.81216
$O12_{poly\text{-}P}$	8.76292	1.54968	7.52738
$O13_{poly\text{-}P}$	9.49269	-0.68770	6.55775
HIS 61			
CB_{His61}	-6.82020	-5.93556	4.89874
$HB1_{His61}$	-7.13749	-6.33548	3.92649
$HB2_{His61}$	-6.56810	-6.80365	5.51732
CG_{His61}	-5.57272	-5.12328	4.68807
$ND1_{His61}$	-5.27160	-4.55384	3.45983
$HD1_{His61}$	-5.92199	-4.47492	2.69166

$CE1_{His61}$	-4.03939	-3.97087	3.54663
$HE1_{His61}$	-3.58045	-3.43365	2.72779
$NE2_{His61}$	-3.51388	-4.14704	4.74326
$HE2_{His61}$	-1.52921	-3.67010	5.19915
$CD2_{His61}$	-4.46658	-4.85067	5.45881
$HD2_{His61}$	-4.31802	-5.11001	6.49538
HIS 220			
CB_{His220}	-2.18375	1.70823	5.57561
$HB1_{His220}$	-1.63312	1.73152	6.52153
$HB2_{His220}$	-1.88324	2.62118	5.06680
CG_{His220}	-1.66352	0.55989	4.77657
$ND1_{His220}$	-0.85286	0.76187	3.66998
$HD1_{His220}$	-0.52739	1.72979	3.26898
$CE1_{His220}$	-0.28172	-0.38585	3.30839
$HE1_{His220}$	0.49497	-0.49983	2.57566
$NE2_{His220}$	-0.72701	-1.34392	4.13053
$CD2_{His220}$	-1.60134	-0.77491	5.04244
$HD2_{His220}$	-2.04737	-1.36244	5.82257
ASP 114			
CB_{Asp114}	2.01819	-5.16902	0.38555
$HB1_{Asp114}$	2.10672	-4.58027	-0.53140
$HB2_{Asp114}$	1.81639	-4.48924	1.21999
CG_{Asp114}	0.80899	-6.11491	0.27608
$OD1_{Asp114}$	0.31789	-6.52974	1.39364
$OD2_{Asp114}$	0.31482	-6.41197	-0.83673

Table S3. Atomic charges on key atoms of the TS localized at B3LYP/MM level of the *substrate-assisted mechanism* in the wild type enzyme.

$O5_{Dha}$	-0.662
$H5_{Dha}$	0.344
$O1_{poly\text{-}P}$	-0.739
$P2_{poly\text{-}P}$	1.107
$O3_{poly\text{-}P}$	-1.044
$O4_{poly\text{-}P}$	-0.651
$O5_{poly\text{-}P}$	-0.616
$P6_{poly\text{-}P}$	1.368
$O7_{poly\text{-}P}$	-0.885
$O8_{poly\text{-}P}$	-0.910

Table S4. Atomic charges on key atoms of the TS localized at B3LYP/MM level of the *substrate-assisted mechanism* in the mutated enzyme.

$O5_{Dha}$	-0.615
$H5_{Dha}$	0.342
$O1_{poly\text{-}P}$	-0.738
$P2_{poly\text{-}P}$	1.102
$O3_{poly\text{-}P}$	-1.049
$O4_{poly\text{-}P}$	-0.656
$O5_{poly\text{-}P}$	-0.618
$P6_{poly\text{-}P}$	1.383
$O7_{poly\text{-}P}$	-0.874
$O8_{poly\text{-}P}$	-0.890

7.5. Mutant Ser105Asp CALB as a new catalyst for epoxide hydrolysis

Is Promiscuous CALB a Good Scaffold for Designing New Epoxidases?

Isabel Bordes, José Recatalá, Katarzyna Swiderek and Vicent Moliner

Published in *Molecules* **2015**, 20, 17789-17806

Attached in Appendix

8. References

(1) Lehninger, A. L.; Bozal Fes, J.; Calvet Prats, F. *Bioquímica: Las bases moleculares de la estructura y función celular*, 2nd Edition, Omega, Barcelona, 2003.
(2) Szarka, A. *Basic biochemistry. University notes for chemical engineers*, Budapest, 2014.
(3) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 7th Edition, W.H.Freeman, New York, 2017.
(4) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. *Chemical Reviews* **2006**, *106*, 3188.
(5) Olsson, M. H. M.; Parson, W. W.; Warshel, A. *Chemical Reviews* **2006**, *106*, 1737.
(6) Carvalho, A. T. P.; Barrozo, A.; Doron, D.; Kilshtain, A. V.; Major, D. T.; Kamerlin, S. C. L. *Journal of Molecular Graphics & Modelling* **2014**, *54*, 62.
(7) Sousa, S. F.; Ramos, M. J.; Lim, C.; Fernandes, P. A. *ACS Catalysis* **2015**, *5*, 5877.
(8) Piazzetta, P.; Marino, T.; Russo, N. *Archives of Biochemistry and Biophysics* **2015**, *582*, 101.
(9) Souza, B. S.; Mora, J. R.; Wanderlind, E. H.; Clementin, R. M.; Gesser, J. C.; Fiedler, H. D.; Nome, F.; Menger, F. M. *Angewandte Chemie-International Edition* **2017**, *56*, 5345.
(10) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *Chemical Society Reviews* **2012**, *41*, 3025.
(11) Tuñón, I.; Moliner, V. *Simulating Enzyme Reactivity. Computational Methods in Enzyme Catalysis*, The Royal Society of Chemistry, Cambridge, 2016.
(12) Winslow, R. L.; Trayanova, N.; Geman, D.; Miller, M. I. *Science Translational Medicine* **2012**, *4(158)*, 158rv11.
(13) Martinez, T. J. *Accounts of Chemical Research* **2017**, *50*, 652.
(14) Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E. *Annual Review of Biophysics* **2012**, *41*, 429.
(15) Esteban, S. *Introducción a la Historia de la Química*, UNED, Madrid, 2011.
(16) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentos de bioquímica: La vida a nivel molecular*, 2nd Edition, Medica paramenicana, Buenos Aires, 2007.
(17) Sumner, J. B. *Journal of Biological Chemistry* **1926**, *69*, 435.
(18) Hirs, C. H. W.; Moore, S.; Stein, W. H. *Journal of Biological Chemistry* **1960**, *235*, 633.
(19) Blake, C. C. F.; Johnson, L. N.; Mair, G. A.; North, A. C. T.; Phillips, D. C.; Sarma, V. R. *Proceedings of the Royal Society Series B-Biological Sciences* **1967**, *167*, 378.
(20) Elliot, W. H. E., Daphne C. *Biochemistry and Molecular Biology*, 3rd Edition, Oxford University Press, United States, 2005.
(21) Mark Berg, J.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 5th Edition, W. H. Freeman, New York, 2002.
(22) Frey, P. A.; Hegeman, A. D. *Enzymatic Reaction Mechanisms*, Oxford University Press, New York, 2007.
(23) Cavaco-Paulo, A.; Gübitz, G. M. *Textile processing with enzymes*, 1st Edition, Woodhead Publishing Limited, Cambridge, 2003.
(24) Sammond, D. W.; Kastelowitz, N.; Himmel, M. E.; Yin, H.; Crowley, M. F.; Bomble, Y. J. *Plos One* **2016**, *11*, e145848.

(25) Alponti, J. S.; Maldonado, R. F.; Ward, R. J. *International Journal of Biological Macromolecules* **2016**, *87*, 522.
(26) Michaelis, L.; Menten, M. L. *Biochemische Zeitschrift* **1913**, *49*, 333.
(27) Fischer, E. *Berichte der deutschen chemischen Gesellschaft* **1894**, *27*, 2985.
(28) Polanyi, M. *Zeitschrift Fur Elektrochemie Und Angewandte Physikalische Chemie* **1921**, *27*, 142.
(29) Pauling, L. *Chemical & Engineering News Archive* **1946**, *24*, 1375.
(30) Martí, S.; Roca, M.; Andres, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertran, J. *Chemical Society Reviews* **2004**, *33*, 98.
(31) Warshel, A. *Proceedings of the National Academy of Sciences of the United States of America* **1978**, *75*, 5250.
(32) Warshel, A. *Journal of Biological Chemistry* **1998**, *273*, 27035.
(33) Villa, J.; Warshel, A. *Journal of Physical Chemistry B* **2001**, *105*, 7887.
(34) Cheek, S.; Ginalski, K.; Zhang, H.; Grishin, N. V. *BMC Structural Biology* **2005**, *5*, 1.
(35) Kenyon, C. P.; Roth, R. L.; van der Westhuyzen, C. W.; Parkinson, C. J. *BMC research notes* **2012**, *5*, 131.
(36) Jacobsen, D. M.; Bao, Z.-Q.; O'Brien, P.; Brooks, C. L., III; Young, M. A. *Journal of the American Chemical Society* **2012**, *134*, 15357.
(37) Roskoski Jr., R. *Pharmacological Research* **2015**, *100*, 1.
(38) Krebs, E. G.; Graves, D. J.; Fischer, E. H. *Journal of Biological Chemistry* **1959**, *234*, 2867.
(39) Walsh, D. A.; Perkins, J. P.; Krebs, E. G. *Journal of Biological Chemistry* **1968**, *243*, 3763.
(40) Eckhart, W.; Hutchinson, M. A.; Hunter, T. *Cell* **1979**, *18*, 925.
(41) Hunter, T.; Sefton, B. M. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1980**, *77*, 1311.
(42) Hunter, T. *Cell* **1987**, *50*, 823.
(43) Barker, W. C.; Dayhoff, M. O. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1982**, *79*, 2836.
(44) Knighton, D. R.; Zheng, J. H.; Teneyck, L. F.; Ashford, V. A.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. *Science* **1991**, *253*, 407.
(45) Knighton, D. R.; Zheng, J. H.; Teneyck, L. F.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. *Science* **1991**, *253*, 414.
(46) Kornev, A. P.; Taylor, S. S. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2010**, *1804*, 440.
(47) Srivastava, A. K.; McDonald, L. R.; Cembran, A.; Kim, J.; Masterson, L. R.; McClendon, C. L.; Taylor, S. S.; Veglia, G. *Structure* **2014**, *22*, 1735.
(48) Pérez-Gallegos, A.; García-Viloca, M.; González-Lafont, A.; Lluch, J. M. *Physical Chemistry Chemical Physics* **2015**, *17*, 3497.
(49) Margutti, S.; Laufer, S. A. *Chemmedchem* **2007**, *2*, 1116.
(50) Kamerlin, S. C. L.; Williams, N. H.; Warshel, A. *Journal of Organic Chemistry* **2008**, *73*, 6960.
(51) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. *Annual Review of Biochemistry* **2011**, *80*, 669.

(52) Smith, G. K.; Ke, Z.; Guo, H.; Hengge, A. C. *Journal of Physical Chemistry B* **2011**, *115*, 13713.
(53) Ptacek, J.; Snyder, M. *Trends in Genetics* **2006**, *22*, 545.
(54) Kinnings, S. L.; Jackson, R. M. *Journal of Chemical Information and Modeling* **2009**, *49*, 318.
(55) Zhang, J.; Yang, P. L.; Gray, N. S. *Nature Reviews Cancer* **2009**, *9*, 28.
(56) Colinge, J.; César-Razquin, A.; Huber, K.; Breitwieser, F. P.; Májek, P.; Superti-Furga, G. *Journal of Proteomics* **2014**, *107*, 113.
(57) Volkamer, A.; Eid, S.; Turk, S.; Jaeger, S.; Rippmann, F.; Fulle, S. *Journal of Chemical Information and Modeling* **2015**, *55*, 538.
(58) Muller, G. *Drug Discovery Today* **2003**, *8*, 681.
(59) Stout, T. J.; Foster, P. G.; Matthews, D. J. *Current pharmaceutical design* **2004**, *10*, 1069.
(60) Cao, D.-S.; Zhou, G.-H.; Liu, S.; Zhang, L.-X.; Xu, Q.-S.; He, M.; Liang, Y.-Z. *Analytica Chimica Acta* **2013**, *792*, 10.
(61) Ryall, K. A.; Shin, J.; Yoo, M.; Hinz, T. K.; Kim, J.; Kang, J.; Heasley, L. E.; Tan, A. C. *Bioinformatics* **2015**, *31*, 3799.
(62) Weinmann, H.; Metternich, R. *Chembiochem* **2005**, *6*, 455.
(63) Bamborough, P. *Expert Opinion on Drug Discovery* **2012**, *7*, 1053.
(64) Bossemeyer, D. *Febs Letters* **1995**, *369*, 57.
(65) Hart, J. C.; Sheppard, D. W.; Hillier, I. H.; Burton, N. A. *Chemical Communications* **1999**, 79.
(66) Milletti, F.; Vulpetti, A. *Journal of Chemical Information and Modeling* **2010**, *50*, 1418.
(67) Kumar, R. In *Nuclear Signaling Pathways and Targeting Transcription in Cancer*; Springer New York: New York, 2014.
(68) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912.
(69) Zurbriggen, A.; Jeckelmann, J.-M.; Christen, S.; Bieniossek, C.; Baumann, U.; Erni, B. *Journal of Biological Chemistry* **2008**, *283*, 35789.
(70) Shi, R.; McDonald, L.; Cui, Q.; Matte, A.; Cygler, M.; Ekiel, I. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 1302.
(71) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826.
(72) Enders, D.; Voith, M.; Lenzen, A. *Angewandte Chemie-International Edition* **2005**, *44*, 1304.
(73) Sánchez-Moreno, I.; Bordes, I.; Castillo, R.; Ruiz-Pernía, J. J.; Moliner, V.; García-Junceda, E. *International Journal of Molecular Sciences* **2015**, *16*, 26073.
(74) García-Alles, L. F.; Siebolo, C.; Nyffeler, T. L.; Flukiger-Bruhwiler, K.; Schneider, P.; Burgi, H. B.; Baumann, U.; Erni, B. *Biochemistry* **2004**, *43*, 13037.
(75) Bachler, C.; Flukiger-Bruhwiler, K.; Schneider, P.; Bahler, P.; Erni, B. *Journal of Biological Chemistry* **2005**, *280*, 18321.
(76) Erni, B.; Siebold, C.; Christen, S.; Srinivas, A.; Oberholzer, A.; Baumann, U. *Cellular and Molecular Life Sciences* **2006**, *63*, 890.

(77) Raynaud, C.; Lee, J.; Sarcabal, P.; Croux, C.; Meynial-Salles, I.; Soucaille, P. *Journal of Bacteriology* **2011**, *193*, 3127.
(78) Shi, R.; McDonald, L.; Cygler, M.; Ekiel, I. *Structure* **2014**, *22*, 478.
(79) Cleland, W. W.; Hengge, A. C. *Chemical Reviews* **2006**, *106*, 3252.
(80) Marcos, E.; Crehuet, R.; Anglada, J. M. *Journal of Chemical Theory and Computation* **2008**, *4*, 49.
(81) Kamerlin, S. C. L. *Journal of Organic Chemistry* **2011**, *76*, 9228.
(82) Kamerlin, S. C.; Sharma, P. K.; Prasad, R. B.; Warshel, A. *Q Rev Biophys* **2013**, *46*, 1.
(83) Wymore, T.; Field, M. J.; Langan, P.; Smith, J. C.; Parks, J. M. *Journal of Physical Chemistry B* **2014**, *118*, 4479.
(84) Kamerlin, S. C. L.; Wilkie, J. *Organic & Biomolecular Chemistry* **2007**, *5*, 2098.
(85) Kamerlin, S. C. L.; Florian, J.; Warshel, A. *Chemphyschem* **2008**, *9*, 1767.
(86) Rosta, E.; Kamerlin, S. C. L.; Warshel, A. *Biochemistry* **2008**, *47*, 3725.
(87) Diaz, N.; Field, M. J. *Journal of the American Chemical Society* **2004**, *126*, 529.
(88) Cheng, Y. H.; Zhang, Y. K.; McCammon, J. A. *Journal of the American Chemical Society* **2005**, *127*, 1553.
(89) Pérez-Gallegos, A.; García-Viloca, M.; González-Lafont, A.; Lluch, J. M. *ACS Catalysis* **2015**, *5*, 4897.
(90) Ojeda-May, P.; Li, Y.; Ovchinnikov, V.; Nam, K. *Journal of the American Chemical Society* **2015**, *137*, 12454.
(91) Hutter, M. C.; Helms, V. *Protein Science* **1999**, *8*, 2728.
(92) De Vivo, M.; Cavalli, A.; Carloni, P.; Recanatini, M. *Chemistry-a European Journal* **2007**, *13*, 8437.
(93) Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864.
(94) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chemical Physics Letters* **1989**, *157*, 200.
(95) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648.
(96) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *The Journal of Chemical Physics* **1982**, *77*, 3654.
(97) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chemical Physics Letters* **1988**, *153*, 503.
(98) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *Journal of Computational Chemistry* **1983**, *4*, 294.
(99) Woon, D. E.; Dunning, T. H. *Journal of Chemical Physics* **1993**, *98*, 1358.
(100) Nam, K.; Cui, Q.; Gao, J.; York, D. M. *Journal of Chemical Theory and Computation* **2007**, *3*, 486.
(101) Huang, M.; Li, X.; Zou, J.-W.; Timson, D. J. *Biochemistry* **2013**, *52*, 4858.
(102) Bastidas, A. C.; Deal, M. S.; Steichen, J. M.; Guo, Y.; Wu, J.; Taylor, S. S. *Journal of the American Chemical Society* **2013**, *135*, 4788.
(103) Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Vu, T.; Alber, T.; Hagan, M. F.; Kern, D. *Nature Structural & Molecular Biology* **2015**, *22*, 124.
(104) Hartree, D. R. *Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 89.

(105) Harrison, C. B.; Schulten, K. *Journal of Chemical Theory and Computation* **2012**, *8*, 2328.
(106) López-Canut, V.; Martí, S.; Bertran, J.; Moliner, V.; Tuñón, I. *Journal of Physical Chemistry B* **2009**, *113*, 7816.
(107) López-Canut, V.; Roca, M.; Bertran, J.; Moliner, V.; Tuñón, I. *Journal of the American Chemical Society* **2010**, *132*, 6955.
(108) López-Canut, V.; Roca, M.; Bertran, J.; Moliner, V.; Tuñón, I. *Journal of the American Chemical Society* **2011**, *133*, 12050.
(109) López-Canut, V.; Ruiz-Pernía, J. J.; Castillo, R.; Moliner, V.; Tuñón, I. *Chemistry-a European Journal* **2012**, *18*, 9612.
(110) Lopata, A.; Jambrina, P. G.; Sharma, P. K.; Brooks, B. R.; Toth, J.; Vertessy, B. G.; Rosta, E. *ACS Catalysis* **2015**, *5*, 3225.
(111) Postma, P. W.; Lengeler, J. W.; Jacobson, G. R. *Microbiological Reviews* **1993**, *57*, 543.
(112) Deutscher, J.; Francke, C.; Postma, P. W. *Microbiology and Molecular Biology Reviews* **2006**, *70*, 939.
(113) Kundig, W.; Roseman, S.; Ghosh, S. *Proceedings of the National Academy of Sciences of the United States of America* **1964**, *52*, 1067.
(114) Saier Jr., M. H. *Journal of Molecular Microbiology and Biotechnology* **2015**, *25*, 73.
(115) Gutknecht, R.; Beutler, R.; García-Alles, L. F.; Baumann, U.; Erni, B. *Embo Journal* **2001**, *20*, 2480.
(116) Siebold, C.; Arnold, I.; García-Alles, L. F.; Baumann, U.; Erni, B. *Journal of Biological Chemistry* **2003**, *278*, 48236.
(117) Palomo, C.; Oiarbide, M.; García, J. M. *Chemical Society Reviews* **2004**, *33*, 65.
(118) Li, C. J. *Chemical Reviews* **2005**, *105*, 3095.
(119) Schmidt, N. G.; Eger, E.; Kroutil, W. *ACS Catalysis* **2016**, *6*, 4286.
(120) Sánchez-Moreno, I.; García-García, J. F.; Bastida, A.; García-Junceda, E. *Chemical Communications* **2004**, *14*, 1634.
(121) Schumperli, M.; Pellaux, R.; Panke, S. *Applied Microbiology and Biotechnology* **2007**, *75*, 33.
(122) Sánchez-Moreno, I.; Iturrate, L.; Doyagueez, E. G.; Antonio Martínez, J.; Fernández-Mayoralas, A.; García-Junceda, E. *Advanced Synthesis & Catalysis* **2009**, *351*, 2967.
(123) Iturrate, L.; Sánchez-Moreno, I.; Oroz-Guinea, I.; Pérez-Gil, J.; García-Junceda, E. *Chemistry-a European Journal* **2010**, *16*, 4018.
(124) Fesko, K.; Gruber-Khadjawi, M. *Chemcatchem* **2013**, *5*, 1248.
(125) Buchholz, K.; Volker, K.; Bornscheuer, U. T. *Biocatalysts and Enzyme Technology*, 2nd Edition, Wiley-VCH Verlag & Co., Weinheim, Germany, 2012.
(126) Wei, M.; Li, Z.; Li, T.; Wu, B.; Liu, Y.; Qu, J.; Li, X.; Li, L.; Cai, L.; Wang, P. G. *ACS Catalysis* **2015**, *5*, 4060.
(127) Yang, J.; Zhu, Y.; Men, Y.; Sun, S.; Zeng, Y.; Zhang, Y.; Sun, Y.; Ma, Y. *Journal of Agricultural and Food Chemistry* **2016**, *64*, 9497.
(128) Schuemperli, M.; Pellaux, R.; Panke, S. *Applied Microbiology and Biotechnology* **2007**, *75*, 33.

(129) Iwamoto, S.; Motomura, K.; Shinoda, Y.; Urata, M.; Kato, J.; Takiguchi, N.; Ohtake, H.; Hirota, R.; Kuroda, A. *Applied and Environmental Microbiology* **2007**, *73*, 5676.
(130) Restiawaty, E.; Iwasa, Y.; Maya, S.; Honda, K.; Omasa, T.; Hirota, R.; Kuroda, A.; Ohtake, H. *Process Biochemistry* **2011**, *46*, 1747.
(131) Kameda, A.; Shiba, T.; Kawazoe, Y.; Satoh, Y.; Ihara, Y.; Munekata, M.; Ishige, K.; Noguchi, T. *Journal of Bioscience and Bioengineering* **2001**, *91*, 557.
(132) Sato, M.; Masuda, Y.; Kirimura, K.; Kino, K. *Journal of Bioscience and Bioengineering* **2007**, *103*, 179.
(133) Brown, M. R. W.; Kornberg, A. *Trends in Biochemical Sciences* **2008**, *33*, 284.
(134) Rao, N. N.; Gomez-Garcia, M. R.; Kornberg, A. *Annual Review of Biochemistry* **2009**, *78*, 605.
(135) Senn, H. M.; Thiel, W. *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations*, Springer Berlin Heidelberg, Berlin, 2007.
(136) Leach, A. R. *Molecular Modelling. Principles and Applications*, 2nd Edition, Pearson Education, Harlow, 2001.
(137) Jensen, F. *Introduction to Computational Chemistry*, 2nd Edition, John Wiley & Sons, Chichester, 2007.
(138) Warshel, A.; Karplus, M. *Journal of the American Chemical Society* **1974**, *96*, 5677.
(139) Warshel, A.; Levitt, M. *Journal of Molecular Biology* **1976**, *103*, 227.
(140) Field, M. J.; Bash, P. A.; Karplus, M. *Journal of Computational Chemistry* **1990**, *11*, 700.
(141) Gao, J. L. *Accounts of Chemical Research* **1996**, *29*, 298.
(142) van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708.
(143) Born, M.; Oppenheimer, R. *Annalen Der Physik* **1927**, *84*, 0457.
(144) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjovoll, M.; Fahmi, A.; Schafer, A.; Lennartz, C. *Journal of Molecular Structure-Theochem* **2003**, *632*, 1.
(145) Senn, H. M.; Thiel, W. *Angewandte Chemie-International Edition* **2009**, *48*, 1198.
(146) Maseras, F.; Morokuma, K. *Journal of Computational Chemistry* **1995**, *16*, 1170.
(147) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *Journal of Physical Chemistry* **1996**, *100*, 19357.
(148) Moreira, C.; Ramos, M. J.; Fernandes, P. A. *Chemistry-A European Journal* **2016**, *22*, 9218.
(149) Pirillo, J.; Mazzone, G.; Russo, N.; Bertinil, L. *Journal of Chemical Information and Modeling* **2017**, *57*, 234.
(150) Field, M. *A Practical Introduction to the Simulation of Molecular Systems*, 2nd Edition, Cambridge University Press, Cambridge, 2007.

(151) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *Journal of Physical Chemistry B* **1998**, *102*, 3586.
(152) Mackerell, A. D.; Feig, M.; Brooks, C. L. *Journal of Computational Chemistry* **2004**, *25*, 1400.
(153) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *Journal of Computational Chemistry* **2010**, *31*, 671.
(154) Jorgensen, W. L.; Tiradorives, J. *Journal of the American Chemical Society* **1988**, *110*, 1657.
(155) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983**, *79*, 926.
(156) Tee, L. S.; Gotoh, S.; Stewart, W. E. *Industrial & Engineering Chemistry Fundamentals* **1966**, *5*, 356.
(157) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *Journal of Physical Chemistry A* **2000**, *104*, 1720.
(158) Hartree, D. R. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 89.
(159) Bertran Rusca, J.; Branchadell Gallo, V.; Moreno Ferrer, M.; Sodupe Roure, M. *Química cuántica. Fundamentos y aplicaciones computacionales*, 2nd Edition, Madrid, 2002.
(160) Florian, J.; Warshel, A. *Journal of Physical Chemistry B* **1998**, *102*, 719.
(161) Liu, Y.; Lopez, X.; York, D. M. *Chemical Communications* **2005**, 3909.
(162) Lopez, X.; Dejaegere, A.; Leclerc, F.; York, D. M.; Karplus, M. *Journal of Physical Chemistry B* **2006**, *110*, 11525.
(163) Duarte, F.; Barrozo, A.; Åqvist, J.; Williams, N. H.; Kamerlin, S. C. L. *Journal of the American Chemical Society* **2016**, *138*, 10664.
(164) Aranda, J.; Roca, M.; López-Canut, V.; Tuñón, I. *Journal of Physical Chemistry B* **2010**, *114*, 8467.
(165) Bras, N. F.; Fernandes, P. A.; Ramos, M. J. *ACS Catalysis* **2014**, *4*, 2587.
(166) Fortino, M.; Marino, T.; Russo, N.; Sicilia, E. *Journal of Molecular Modeling* **2016**, *22*, 287.
(167) Medina, F. E.; Neves, R. P. P.; Ramos, M. J.; Fernandes, P. A. *Physical Chemistry Chemical Physics* **2017**, *19*, 347.
(168) Mora, J. R.; Kirby, A. J.; Nome, F. *The Journal of Organic Chemistry* **2012**, *77*, 7061.
(169) Li, Q.; Fan, S.; Li, X.; Jin, Y.; He, W.; Zhou, J.; Cen, S.; Yang, Z. *Scientific Reports* **2016**, *6*, 38088.
(170) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *Journal of the American Chemical Society* **1985**, *107*, 3902.
(171) Stewart, J. J. P. *Journal of Computational Chemistry* **1989**, *10*, 209.
(172) Zhao, Y.; Truhlar, D. G. *Theoretical Chemistry Accounts* **2008**, *120*, 215.

(173) Nam, K.; Gaot, J.; York, D. M. *Journal of the American Chemical Society* **2008**, *130*, 4680.
(174) López-Canut, V.; Ruiz-Pernía, J. J.; Tuñón, I.; Ferrer, S.; Moliner, V. *Journal of Chemical Theory and Computation* **2009**, *5*, 439.
(175) Brandt, W.; Dessoy, M. A.; Fulhorst, M.; Gao, W. Y.; Zenk, M. H.; Wessjohann, L. A. *Chembiochem* **2004**, *5*, 311.
(176) Hand, C. E.; Honek, J. F. *Bioorganic & Medicinal Chemistry Letters* **2007**, *17*, 183.
(177) Plotnikov, N. V.; Prasad, B. R.; Chakrabarty, S.; Chu, Z. T.; Warshel, A. *The Journal of Physical Chemistry B* **2013**, *117*, 12807.
(178) Stewart, J. J. P. *Journal of Molecular Modeling* **2007**, *13*, 1173.
(179) Nguyen, K. A.; Rossi, I.; Truhlar, D. G. *Journal of Chemical Physics* **1995**, *103*, 5522.
(180) Corchado, J. C.; Coitino, E. L.; Chuang, Y. Y.; Fast, P. L.; Truhlar, D. G. *Journal of Physical Chemistry A* **1998**, *102*, 2424.
(181) Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G. *Journal of Physical Chemistry A* **1999**, *103*, 1140.
(182) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. *Journal of Physical Chemistry B* **2004**, *108*, 8427.
(183) Roca, M.; Moliner, V.; Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I. *Journal of Physical Chemistry A* **2006**, *110*, 503.
(184) Renka, R. J. *Siam Journal on Scientific and Statistical Computing* **1987**, *8*, 393.
(185) Renka, R. J. *Acm Transactions on Mathematical Software* **1993**, *19*, 81.
(186) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S. *Journal of Physical Chemistry B* **2006**, *110*, 17663.
(187) Tuñón, I.; Silla, E. *Química Molecular Estadística. Termodinámica Estadística para químicos y bioquímicos*, Sintesis, Madrid, 2008.
(188) Havlas, Z.; Zahradnik, R. *International Journal of Quantum Chemistry* **1984**, *26*, 607.
(189) González, C.; Schlegel, H. B. *Journal of Chemical Physics* **1989**, *90*, 2154.
(190) Fukui, K. *Journal of Physical Chemistry* **1970**, *74*, 4161.
(191) Fukui, K. *Accounts of Chemical Research* **1981**, *14*, 363.
(192) Fukui, K. *Recueil Des Travaux Chimiques Des Pays-Bas-Journal of the Royal Netherlands Chemical Society* **1979**, *98*, 75.
(193) Melissas, V. S.; Truhlar, D. G.; Garrett, B. C. *Journal of Chemical Physics* **1992**, *96*, 5758.
(194) Verlet, L. *Physical Review* **1967**, *159*, 98.
(195) Nose, S. *Journal of Chemical Physics* **1984**, *81*, 511.
(196) Hoover, W. G. *Physical Review A* **1985**, *31*, 1695.
(197) Torrie, G. M.; Valleau, J. P. *Journal of Computational Physics* **1977**, *23*, 187.
(198) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *Journal of Computational Chemistry* **1992**, *13*, 1011.
(199) Eyring, H. *Journal of Chemical Physics* **1935**, *3*, 107.
(200) Evans, M. G.; Polanyi, M. *Transactions of the Faraday Society* **1935**, *31*, 0875.
(201) Truhlar, D. G. *Journal of Physical Organic Chemistry* **2010**, *23*, 660.

(202) Truhlar, D. G. *Archives of Biochemistry and Biophysics* **2015**, *582*, 10.
(203) Truhlar, D. G.; Gao, J.; Alhambra, C.; García-Viloca, M.; Corchado, J.; Sánchez, M. L.; Villa, J. *Acc Chem Res* **2002**, *35*, 341.
(204) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. *The Journal of Physical Chemistry* **1996**, *100*, 12771.
(205) Nam, K.; Prat-Resina, X.; García-Viloca, M.; Devi-Kesavan, L. S.; Gao, J. *J Am Chem Soc* **2004**, *126*, 1369.
(206) García-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186.
(207) Truhlar, D. G.; Gao, J.; García-Viloca, M.; Alhambra, C.; Corchado, J.; Luz Sánchez, M.; Poulsen, T. D. *International Journal of Quantum Chemistry* **2004**, *100*, 1136.
(208) Kamerlin, S. C. L.; Warshel, A. *Proteins-Structure Function and Bioinformatics* **2010**, *78*, 1339.
(209) Luk, L. Y. P.; Ruiz-Pernía, J. J.; Dawson, W. M.; Roca, M.; Joel Loveridge, E.; Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J.; Tuñón, I.; Moliner, V.; Allemann, R. K. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 16344.
(210) García-Meseguer, R.; Martí, S.; Ruiz-Pernía, J. J.; Moliner, V.; Tuñón, I. *Nature Chemistry* **2013**, *5*, 566.
(211) Tuñón, I.; Laage, D.; Hynes, J. T. *Archives of Biochemistry and Biophysics* **2015**, *582*, 42.
(212) Klinman, J. P. *Accounts of Chemical Research* **2015**, *48*, 449.
(213) Warshel, A.; Bora, R. P. *Journal of Chemical Physics* **2016**, *144*, 180901.
(214) Basner, J. E.; Schwartz, S. D. *Journal of the American Chemical Society* **2005**, *127*, 13822.
(215) Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T. *Nature* **2009**, *462*, 669.
(216) Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J. *Nature Chemistry* **2012**, *4*, 169.
(217) Layfield, J. P.; Hammes-Schiffer, S. *Chemical Reviews* **2014**, *114*, 3466.
(218) Dzierlenga, M. W.; Antoniou, D.; Schwartz, S. D. *Journal of Physical Chemistry Letters* **2015**, *6*, 1177.
(219) Sutcliffe, M. J.; Scrutton, N. S. *Trends in Biochemical Sciences* **2000**, *25*, 405.
(220) Dutton, P. L.; Munro, A. W.; Scrutton, N. S.; Sutcliffe, M. J. *Philosophical Transactions of the Royal Society B-Biological Sciences* **2006**, *361*, 1293.
(221) Hwang, J. K.; Warshel, A. *Journal of the American Chemical Society* **1996**, *118*, 11745.

9. Appendix

Poster communications:

- *Theoretical study of molecular mechanism of alpha, beta-epoxyesters.* Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía, Florenci V. González and Vicent Moliner. Ninth Triennial Congress of the World Association of Theoretical and Computational Chemists (WATOC 2011), Santiago de Compostela, 2011.

- *Electron transference pathways in trypsin-solubilized bovine liver cytochrome b*. Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía, Aurélien de la Lande and Vicent Moliner. 8th Congress on Electronic Structure: Principles and Applications (ESPA 2012), Barcelona, 2012.

- *Effects of possible mutations in dihydroxyacetone kinase using inorganic polyphosphate as phosphate donor.* Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía, Vicent Moliner and Eduardo García-Junceda. XXXIX International Conference; Theoretical Chemists of Latin Expression (QUITEL 2013), Granada, 2013.

- *A QM/MM Study of Phosphate Transfer to Dihydroxyacetone in Aqueous Solution.* Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía and Vicent Moliner**.** 9th Congress on Electronic Structure: Principles and Applications (ESPA 2014), Badajoz, 2014.

- *A QM/MM study of Phosphate Transfer to Dihydroxyacetone in Dihydroxyacetone Kinase from Escherichia Coli.* Isabel Bordes, Raquel Castillo, José Javier Ruiz-Pernía and Vicent Moliner**.** 7th International Theoretical Biophysics Symposium (TheoBio 2015), Cagliari (Italia), 2015.

- *Substrate-assisted versus Asp-assisted phosphoryl transfer mechanism in dihydroxyacetone kinase*. Isabel Bordes, Raquel Castillo and Vicent Moliner. 10th Congress on Electronic Structure: Principles and Applications (ESPA 2016), Castellón, 2016.

Other publications:

Is Promiscuous CALB a Good Scaffold for Designing New Epoxidases? Isabel Bordes, José Recatalá, Katarzyna Swiderek and Vicent Moliner. Molecules **2015**, 20, 17789-17806.

Molecules **2015**, *20*, 17789-17806; doi:10.3390/molecules201017789

OPEN ACCESS

molecules

ISSN 1420-3049

www.mdpi.com/journal/molecules

Article

Is Promiscuous CALB a Good Scaffold for Designing New Epoxidases?

Isabel Bordes [1], José Recatalá [1], Katarzyna Świderek [1,2,*] and Vicent Moliner [1,*]

[1] Departament de Química Física i Analítica, Universitat Jaume I, Castellón 12071, Spain; E-Mails: bordes@uji.es (I.B.); al189587@alumail.uji.es (J.R.)

[2] Institute of Applied Radiation Chemistry, Lodz University of Technology, Lodz 90-924, Poland

* Authors to whom correspondence should be addressed; E-Mails: swiderek@uji.es (K.Ś.); moliner@uji.es (V.M.); Tel.: +34-964-728-084 (V.M.); +34-964-728-070 (K.Ś.); Fax: +34-964-728-066 (K.Ś. & V.M.).

Academic Editors: Diego A. Alonso and Isidro M. Pastor

Received: 7 August 2015 / Accepted: 11 September 2015 / Published: 25 September 2015

Abstract: *Candida Antarctica* lipase B (CALB) is a well-known enzyme, especially because of its promiscuous activity. Due to its properties, CALB was widely used as a benchmark for designing new catalysts for important organic reactions. The active site of CALB is very similar to that of soluble epoxide hydrolase (sEH) formed by a nucleophile-histidine-acid catalytic triad and an oxyanion hole typical for molecular structures derived from processes of α/β hydrolases. In this work we are exploring these similarities and proposing a Ser105Asp variant of CALB as a new catalyst for epoxide hydrolysis. In particular, the hydrolysis of the trans-diphenylpropene oxide (t-DPPO) is studied by means of quantum cluster models mimicking the active site of both enzymes. Our results, based on semi-empirical and DFT calculations, suggest that mutant Ser105Asp CALB is a good protein scaffold to be used for the bio-synthesis of chiral compounds.

Keywords: *Candida antarctica* lipase B; CALB; epoxide hydrolase; sEH; reaction mechanism; trans-diphenylpropene oxide; enzyme promiscuity; catalysis; quantum cluster models

Molecules **2015**, *20* **17790**

1. Introduction

Epoxides are important molecules for producing chiral compounds. Because of their chemical versatility and ability to react readily with halides, carbon, nitrogen, oxygen, or sulfur nucleophiles, epoxides became crucial intermediate products in organic synthesis [1]. Epoxides are three-membered cyclic ethers that have specific reactivity patterns characterized by their highly polarized oxygen–carbon bonds in addition to a highly strained ring [2]. In recent years enormous efforts have been done into developing methodologies for preparing enantio-pure forms of epoxides. This purpose has been achieved by applying new chemical [3] and biocatalytic [4] procedures. Biocatalytic conversion of epoxides can proceed by, for instance, conjugation of thiol cofactors, nucleophilic ring opening, or hydrolysis. The last process is, in fact, the subject of the present work.

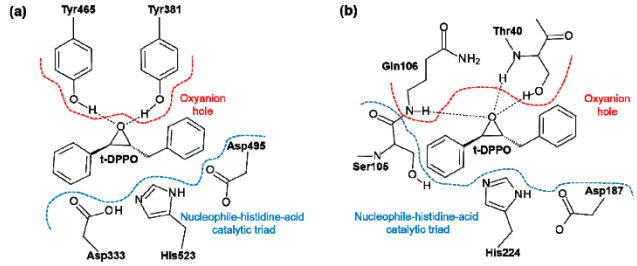
Herein we focus on the epoxide hydrolases (EHs, E.C.3.3.2.3) that serve as a catalyst to transform epoxide into the corresponding 1,2-diol by addition of a water molecule. In fact, mammalian EHs were widely studied mostly because of their biological functions. The main roles of EHs in mammalian organisms are detoxification, catabolism, and regulation of signaling molecules. However, EHs have mainly garnered interest because of their potential applications in chiral chemistry [1,5].

Soluble epoxide hydrolase (sEH) hydrolyses a broad range of substrates such as gem-di-, trans-di-, cis-di, tri-, and tetra-substituted epoxides [6]. However, for the typical *in vitro* purposes, the substrate is trans-diphenylpropene oxide (t-DPPO), and thus it will be used as a substrate in the present study. Recently, it was found that sEH plays an important role in the regulation of blood pressure and inflammation [7–13], which also makes it a good target for designing drugs to be used in the treatment of several diseases, including several aspects of cardiovascular diseases such as inflammation, hypertension, cardiac hypertrophy, and atherosclerosis, or kidney failure [9,13–16].

sEH belongs to the α/β hydrolase fold family of proteins, which are characterized by a Nucleophile-His-Acid catalytic triad evolved to efficiently operate on substrates with different chemical composition or physicochemical properties and in various biological contexts [17]. In the case of human sEH, this triad is formed by Asp333, His523, and Asp495, as presented in Scheme 1a. Moreover, an "oxyanion hole" is formed in the active sites by two tyrosine residues, Tyr-381 and Tyr-465, which are assumed to stabilize the tetrahedral intermediate by protonation or hydrogen bonding interaction of the oxygen atom of the epoxide [18]. The importance of these two residues for the catalytic process was proven experimentally when a 90% decrease in sEH activity was observed when one of the tyrosine residues was mutated to phenylalanine [19].

The structure of the sEH active site is very similar to other enzyme from the α/β hydrolases family named *Candida antarctica* lipase B (CALB), widely studied because of its high catalytic promiscuity [20]. The mechanism of primary reaction of CALB, which is the hydrolysis of ester bonds, was recently studied experimentally [21] and theoretically described [22], providing interesting insights into the functions of residues in the active site. It has been suggested that CALB can be used as an efficient catalyst of biotransformation typical for a carboxylic acid like esterase, thioesterase, peptidase, dehalogenase, epoxide hydrolase, or halo peroxidase, or having the ability to cleave and form C–C bonds [17,23,24]. Moreover, CALB has been also used as a catalyst in ring-opening polymerization reactions [25]. Interestingly, it was shown that both wild-type and Ser105Ala-mutated variants of CALB are able to catalyze direct epoxidation of an α,β-unsaturated aldehyde with hydrogen peroxide [26]. In the present

work, we want to find out if CALB could be used as a scaffold for new catalysts to serve as an epoxide hydrolase.

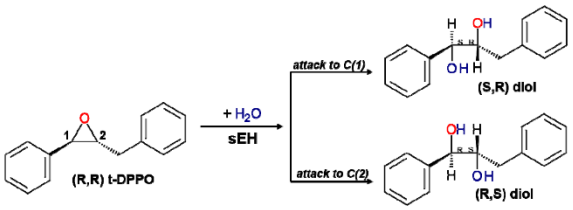


Scheme 1. Active site of (**a**) sEH and (**b**) CALB enzymes together with bound (*R*,*R*)-*trans*-diphenylpropene oxide (t-DPPO).

The active site of CALB contains a Nucleophile-His-Acid catalytic triad formed by Ser105, His224, and Asp187. As in sEH, there is an oxyanion hole in the active site of CALB formed by Thr40 and Gln106 residues. The similarities between the sEH and CALB active sites are indicated on Scheme 1.

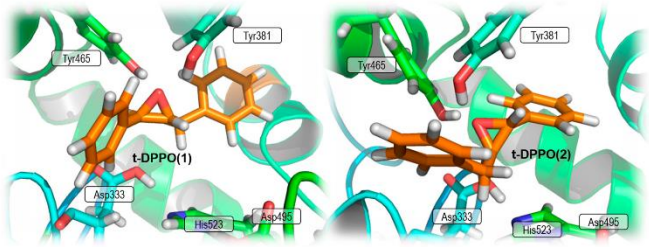
It is generally assumed that the mechanism of hydrolysis of epoxides catalyzed by sEH to the corresponding diols is known. It was demonstrated using experimental techniques that the epoxides can be opened by direct attack of a nucleophile on the epoxide ring or via an intermediate in which a covalent link between the enzyme and the substrate is formed. This variant of the mechanism, in which the nucleophile Asp333 was involved in the reaction, was recently confirmed by theoretical studies performed by Mulholland and co-workers [27]. However, an important question about the origin of the hydrogen that binds to the oxygen of the epoxide after formation of the oxyanion intermediate remains unanswered. Some possible mechanisms have already been proposed, including proton transfer from one of the tyrosine residues that forms the oxyanion hole [28], or direct transfer of a proton from His523 [29]. Nevertheless, these proposals were never explored by theoretical studies. Thus, one of the targets of this work is to dispel these doubts. Recently, a proposal of proton diffusion to the solvent was suggested from a computational, structural, and kinetic study on potato epoxide hydrolase [30].

In the case of hydrolysis of t-DPPO, attacks on both the C1 or C2 carbon atom of epoxide are possible, resulting in two different enantiomers, as presented in Scheme 2.



Scheme 2. Possible products of hydrolysis of (*R*,*R*)-trans-diphenylpropene (t-DPPO).

However, sEH shows regioselectivity by preferentially attacking the reactive C1 carbon atom rather than C2 with a 97:3 selectivity ratio [31]. As was observed in previous theoretical studies [27], this tendency is observed in both possible orientations of substrate t-DPPO(1) and t-DPPO(2), which are related to the position of the substrate relative to the arrangement of residues in the active site of the enzyme, as presented in Figure 1.



(a) **(b)**

Figure 1. Active site of sEH with bound substrate (**a**) t-DPPO(1) and (**b**) t-DPPO(2) orientation.

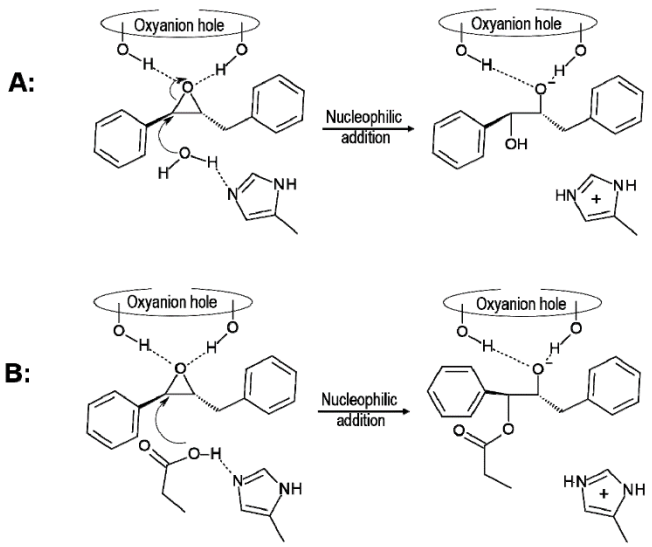
In the present work we are proposing a new catalyst for epoxide hydrolysis reaction based on the structure of the active site of CALB. In order to examine our proposal, the mechanism of the hydrolysis of (*R*,*R*)-t-DPPO in active sites of sEH and a variant of CALB is described using quantum cluster theoretical models.

2. Results and Discussion

2.1. t-DPPO Hydrolysis Catalyzed by sEH

The preparations of the theoretical models require knowledge about the protonation state of titratable residues. Nowadays, fast empirical prediction of pKa algorithm (PROPKA software) [32–35] makes the analysis of protonation states for these residues computationally available. The protonation state of some residues, especially those forming the catalytic triad, seems to be still an open question of debate [29,30,36,37]. In the case of sEH, it was previously observed that protonation of His523 meaningfully influences the rate constant of nucleophilic attack, and it was assumed that this residue has to be protonated in order to reduce the barrier of this step [27]. However, our analysis of pKa results shows that, after substrate binding, the His523 is surely not protonated at pH for sEH activity, which is 7.0–7.5 [6]. This result is in agreement with recent studies on potato sEH [30] and previous studies of Himo and co-workers on human sEH [36,37]. Moreover, it is observed that Asp333, which plays the role of the nucleophile in the studied reaction, has a high pKa value (over 10) (see Table S1 in the Supplementary Materials), which indicates that at the beginning of the hydrolysis it must be protonated, a proposal not suggested in previous studies. Our results indicate that Asp333 has to be activated by transferring the proton and thus acting as the nucleophile that attacks the carbon of epoxide. This step is then part of the full reaction path. Nevertheless, the lack of proton on His523 opens another possible mechanism of this reaction in which, as previously suggested [38], instead of Asp333 a water molecule

can be activated by His523 and subsequently it can attack one of the carbon atoms of the epoxide ring. Thus, both scenarios, (A) attack of water and (B) attack of Asp333, will be taken into consideration in the present study, as depicted in Scheme 3.



Scheme 3. Possible mechanism of nucleophilic attack to the carbon C1 of t-DPPO(1) epoxide by (**A**) a water molecule and (**B**) Asp333.

In the case of mechanism A, the nucleophilic attack can take place concurrently with the proton transfer from a water molecule to Nε-His523. This possibility was investigated by exploring the corresponding potential energy surface (PES), where the antisymmetric combination of distances defined between the oxygen of the water molecule and the carbon atom of epoxide, and between this carbon atom and the oxygen atom of epoxide ($d(O^{wat}–C1)$-$d(O^{epox}–C1)$) was used as one of the distinguished reaction coordinates. The other one was defined as the antisymmetric combination of distances defining the position of the transferred hydrogen atom between the oxygen atoms of the water molecule and the nitrogen atom in position ε of His523 ($d(O^{wat}–H^{wat})$-$d(Nε–H^{wat})$).

In order to explore the second possibility of nucleophilic addition, the nucleophilic attack to C1 of t-DPPO(1) by Asp333, a similar PES was explored. In this case, the antisymmetric combination of distances between oxygen from Asp333 and the carbon of epoxide and between this carbon atom and the oxygen of epoxide ($d(O^{Asp333}–C1)$-$d(O^{epox}–C1)$) was used as one of the coordinates, while the antisymmetric combination of distances describing the transfer of hydrogen atom from oxygen of Asp333 to the nitrogen in position ε of His523 ($d(O^{Asp333}–H^{Asp333})$-$d(Nε–H^{Asp333})$) were controlled as the second reaction coordinate. The resulting PESs are shown in Figure 2. In mechanism A, the nucleophile attack of water occurs simultaneously with proton transfer to Nε-His523. The estimated potential

energy barrier for this step is 38.1 kcal·mol^{-1} at AM1 level and 28.5 kcal·mol^{-1} after correction at M06-2X level. On the other side, mechanism B takes place through a step-wise process whose rate-determining step, the attack of the deprotonated oxygen atom of Asp333 to the carbon atom of the epoxide ring, presents an energy barrier of 39.5 kcal·mol^{-1} at AM1 level but is dramatically reduced to 20.3 kcal·mol^{-1} when recomputed at M06-2X/6-31+G(d,p) level. Thus, our simulations predict mechanism B as the kinetically favorable mechanism and thus we will focus on the next steps of this mechanism. This conclusion is in agreement with experimental studies based on solvent KIEs that confirmed the role of Asp333 residue as the attacking nucleophile since the heavy ^{18}O oxygen was incorporated into the Asp333 of sEH [36]. Moreover, kinetic measurements done for Asp333Ser mutated sEH resulted in a total loss of activity, again indicating the participation of this residue in the hydrolysis of epoxides [39]. Moreover, the obtained value for the energy barrier is in good agreement with the free energy barriers deduced from experimental kinetic measurements at 300 K using the transition state theory (rate constant between 0.09 and 0.75 s^{-1}) [40] for the nucleophilic addition step of hydrolysis of chalconeoxide by sEH. This species can be considered very closely related to the herein-studied t-DPPO substrates. Interestingly, previous theoretical studies for this step resulted in underestimated values of energy barriers equal to 7.8 kcal·mol^{-1} for β-methyl-styrene oxide [32–35] or 9.7 kcal·mol^{-1} for t-DPPO epoxide [27]. In our opinion, this difference comes from the assumption that the reaction pathway begins with negatively charged Asp333, and thus the first step of proton transfer from Asp333 to Nε-His523 was excluded from proposed mechanisms. Keeping in mind these promising results, the study was repeated for the nucleophilic attack of Asp333 to epoxide carbon C1 with the (2)-DPPO orientation, and on the C2 carbon atom of epoxide in the two different orientations of substrate. The results are summarized in Table 1.

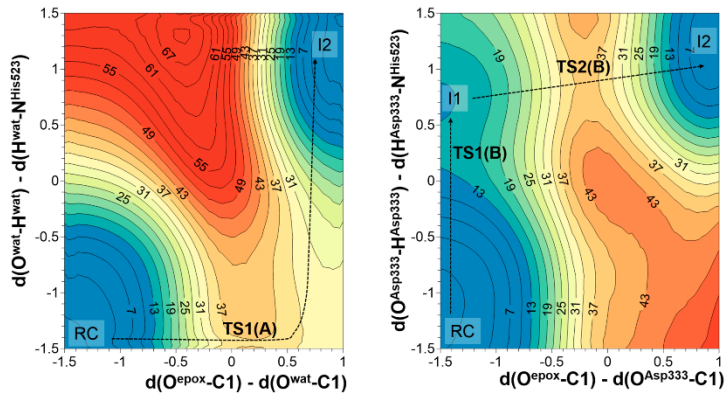


Figure 2. AM1 PES of nucleophilic attack to C1 of t-DPPO(1) by (**a**) a water molecule; and (**b**) Asp333. Distances are in Å and values on iso-energetic lines in kcal·mol^{-1}.

Molecules **2015**, *20* **17795**

Table 1. Relative potential energies obtained at AM1 level related to reactant complex (RC) for nucleophilic attack of Asp333 to epoxide carbons C1 or C2 in two different orientations of substrate in the active site of sEH. Values in parentheses correspond to relative energies of TS2 to I1. Results are given in kcal·mol^{-1}.

	(1)-DPPO				(2)-DPPO			
	C1		C2		C1		C2	
	Cluster	QM/MM [a]	Cluster	QM/MM [a]	Cluster	QM/MM [a]	Cluster	QM/MM [a]
TS1	14.0	-	14.0	-	16.3	-	16.3	-
I1	11.4(0.0)	0.0	11.4	0.0	12.7	0.0	12.7	0.0
TS2	39.9(28.5)	23.5	41.5(30.1)	24.5	42.5(29.8)	18.4	40.2(27.5)	27.6

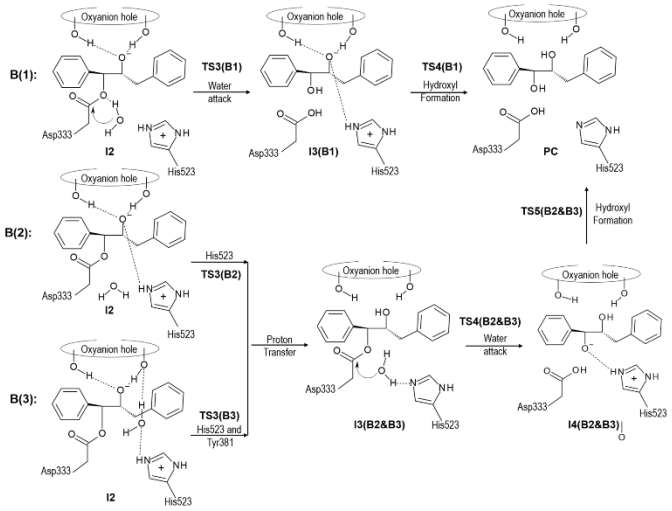
[a]: theoretical values obtained at AM1/MM level by Mulholland and co-workers [27].

Based on the obtained energy barriers for the nucleophilic attack of Asp333 on either the C1 or C2 of epoxide based on our cluster models of the sEH active site, it is difficult to reach conclusions about its enantioselectivity. Comparing our energy barriers presented in Table 1 with the QM/MM theoretical data previously published by Mulholland and co-workers [27], it seems that in the case of t-DPPO(1), some agreement can be found showing that an attack on the C1 carbon of epoxide is more favorable. However, in case of the second orientation t-DPPO(2) we observe opposite tendency. Nevertheless, the differences of energy barriers are too small and thus it is impossible to obtain clear conclusions. It seems that the selectivity of sEH is influenced not only by the shape of the active site itself but by the relation to the global influence of the rest of the protein, probably including the binding effect. Thus, to explore this feature of sEH, QM/MM studies are required. Similar conclusions can be derived from comparison between both orientations of the substrate.

After the first step of the reaction was studied, in which transformation of epoxide to oxyanion tetrahedral intermediate occurs, we focused on the second part of the mechanism, which is still an unsolved problem. The relevance of this step is evident since it was shown by experimental measurements that this step of the reaction, in which the oxyanion intermediate is transformed into a diol, seems to be the rate-limiting step.

According to the obtained results after the nucleophilic addition, the oxyanion intermediate is formed and the protonation state of His523 is changed (from neutral in reactants to positively charged form in the intermediate). The presence of a hydrogen atom in Nε-position does not allow His523 to activate a water molecule in order to attack the carbon of Asp333 and thus alternative scenarios, depicted in Scheme 4, must be considered. In the first one, mechanism B1, we assumed that water directly attacks the carbon atom of Asp333 with a simultaneous proton transfer to the oxygen of the new ether bridge formed between Asp333 and the C1 carbon atom of the epoxide. After this attack, a final proton transfers from Nε-histidine to the oxygen of epoxide occurs and hydroxyl is formed. In the second possibility, mechanism B2 and B3, the reverse order is explored. In mechanism B2, it is assumed that the proton from Nε-His523 is transferred to the oxygen of epoxide, resulting in neutral histidine, which will then be prepared for the last step of the di-hydroxyl formation process. During the second step, a water molecule is activated by transferring its proton to His523 concomitantly with the attack on the carbon of Asp333. Subsequently, the last proton transfers from Nε-His523 to oxygen bound to the C1 carbon of epoxide, resulting in the expected product. As observed, the difference between mechanism

B2 and B3 is the result of different proposals of the origin of hydrogen that is transferred to the negatively charged oxygen atom of epoxide formed in the first step of the process. Thus, in mechanism B3 the proton is transferred from His523 through a water molecule and Tyr381.



Scheme 4. Possible scenarios of hydroxyl formation from oxyanion tetrahedral intermediate of t-DPPO(1) in active site of sEH.

The results obtained for mechanism B1 show that the nucleophilic attack of the water molecule on the carbonyl carbon atom of Asp333 takes place through a TS with an energy, related to the reactant complex, of 64.8 kcal·mol $^{-1}$ at AM1 level and 57.3 kcal·mol $^{-1}$ when corrected at M06-2X/6-31G+(d,p) level (see Figure 3a). In the case of the other two mechanisms, the proton transfer from His523 to the negatively charged oxygen atom of epoxide, either directly (mechanism B2) or through a water molecule and Tyr381 (mechanism B3), takes place with lower potential energy barriers: 4 kcal·mol $^{-1}$ in the case of mechanism B2 (see Figure 3b), and through a stepwise mechanism of 2 and 4 kcal·mol^{-1} in the case of mechanism B3 (see Figure 3c). Afterwards, the nucleophilic attack of the water molecule and the proton transfer to His523 takes place. The corresponding PES, presented in Figure 3d, indicates that, as in mechanism B1, water attack occurs simultaneously with the proton transfer to Nε-His523 (Figure 3a) with a TS that would be 17.2 kcal·mol^{-1} over RC. Interestingly, it seems that the origin of the difference in energy barriers between mechanism B1 and B2 & B3 for this step lies not necessarily in the difference between the proton acceptors, but rather in the geometry of the transition states (TSs). Finally, the hydroxyl formation takes place in a single step in both mechanisms B1 or mechanism B2 & B3, with a potential energy barrier of 2.2 kcal·mol $^{-1}$. All in all, the presented results allow us to conclude that hydroxyl formation from an oxyanion intermediate would preferentially occur through mechanism B2 or B3. The amount of tyrosinate formed during the nucleophilic attack was experimentally estimated to

be 0.4 per enzyme molecule [41], indicating only partial ionization of the residues of the oxyanion hole and, implicitly, the relative relevance of mechanism B2 in the reaction mechanism.

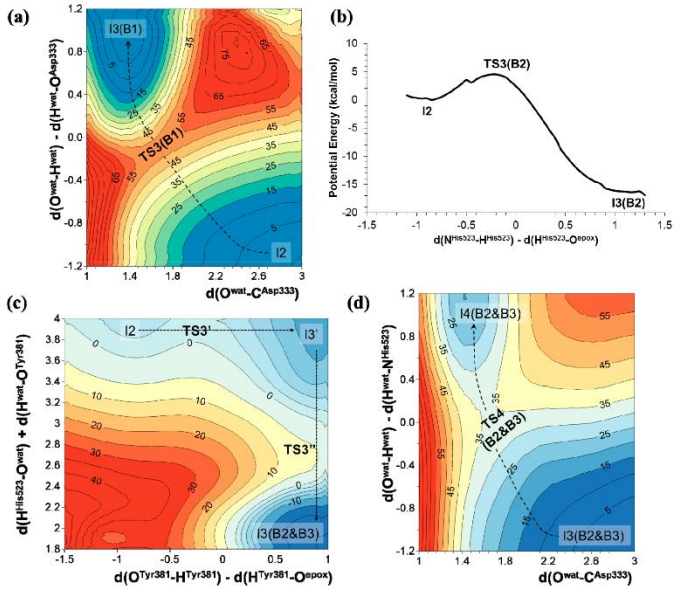


Figure 3. AM1 PESs for hydroxyl formation from oxyanion tetrahedral intermediate of t-DPPO(1) through different possible mechanisms: (**a**) nucleophilic attack of a water molecule on the carbonyl carbon atom of Asp333; (**b**) direct proton transfer from His523 to the negatively charged oxygen atom of epoxide; and (**c**) proton transfer from His523 to the negatively charged oxygen atom of epoxide through a water molecule and Tyr381; (**d**) AM1 PES corresponding to the nucleophilic attack of a water molecule and the proton transfer to His523 from hydroxyl intermediate. All distances are in Å and energies in kcal·mol^{-1}.

2.2. t-DPPO Hydrolysis Catalyzed by CALB

As mentioned in the Introduction, the active site of CALB presents noticeable similarities with the active site of sEH. However, in the wild-type CALB the role of the nucleophile in its catalytic triad is played by a serine (Ser105) instead of an aspartate (Asp333) residue. In fact, as experimentally shown, the Asp333Ser variant of sEH loses its capability to catalyze the epoxide hydrolysis. Thus, it can be concluded that Ser105Asp CALB should perfectly mimic the active site of wild-type sEH. Apart from differences in the rest of the proteins' scaffolds, the main difference between the active sites of sEH and Ser105Asp CALB is the residues involved in the formation of the oxyanion hole. Then, we can predict that the reaction in the latter should proceed through a path similar to mechanism B2 described

for sEH. As we have already demonstrated, the role of the oxyanion hole in sEH is to stabilize the negatively charged oxygen of the intermediate, because the other possible role being a proton donor has already been excluded or does not present any energetic advantage (mechanism B3). Thus, the epoxide hydrolysis in the active site of wild-type CALB and Ser105Asp variant is studied assuming a reaction path similar to the most favorable one observed in sEH, mechanism B2.

The resulting PESs of the nucleophilic attack to C1 carbon atom of t-DPPO(1) in wild-type and Ser105Asp CALB are shown in Figure 4. The corresponding PESs for the attack on the C2 carbon atom, as well as the PESs for the attack to C1 and C2 with the t-DPPO(2) orientation, are reported in Figures S1 and S2 of the Supplementary Materials, showing the same trend as that observed in Figure 4a,b. The analysis of PESs presented in Figure 4 shows a meaningful difference between the wild-type and Ser105Asp variants of CALB. First of all, in the case of wild-type CALB, the Ser105 attack to epoxide and the proton transfer to Nε-His523 take place in a concerted way while, in mutated CALB, attack of Asp105 occurs after proton transfer in a step-wise manner. Moreover, as presented in Table 2, the results show how, in agreement with the results obtained for the epoxide hydrolysis in the active site of sEH, the small differences in the energy barriers between different orientations of the substrate and the two possible nucleophilic attacks to C1 or C2 do not allow us to pinpoint any preference for the formation of one conformer. Consequently, it appears that quantum cluster models present limitations for subtle differences between the energy of conformers in this system. It is important to point out that, considering the good agreement between the trends obtained at the AM1 and M06-2X level of theory for the reaction studied in the sEH, and taking into account the high computational cost of repeating the calculations for this system at a higher level of theory, we report values obtained at the semi-empirical level. It is, then, reasonable to predict that barriers would decrease dramatically if computed at the DFT level of theory.

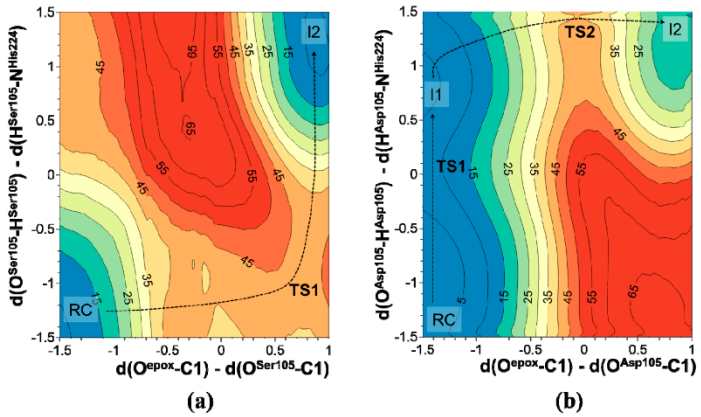


Figure 4. PES of the nucleophilic attack to C1 carbon atom of t-DPPO(1) by (**a**) Ser105 in wild-type CALB and (**b**) Asp105 in Ser105Asp CALB.

Table 2. Relative potential energies obtained at AM1 level related to reactant complex (RC) for the nucleophilic attack of Asp105 in Ser105Asp CALB on epoxide carbons C1 or C2 in two different orientations of substrate. Values are given in kcal·mol $^{-1}$.

	(1)-DPPO		(2)-DPPO	
	C1	**C2**	**C1**	**C2**
RC	0.0	0.0	0.0	0.0
TS1	9.8	9.8	7.0	7.0
I1	2.6	2.6	−0.5	−0.5
I1	−0.8	0.3	2.0	−1.8
TS2	39.5	43.0	44.8	42.1

In order to compare the pathway of the reaction catalyzed by she or by the Ser105Asp variant of CALB, the rest of the steps from INT3 to PC were explored following the same strategy as the one employed in the study of mechanism B2 & B3 of the reaction catalyzed by sEH. The full energy profiles are summarized in Figure 5, while a ball and stick representation of the stationary point structures is shown in Figures 6 and 7, where key inter-atomic distances and angles are reported. The first conclusion that can be derived from an analysis of Figure 5 is that both reaction paths are quite similar, regarding the number of steps and the kind of transformations taking place in each step. Nevertheless, while the rate-limiting step for hydrolysis of t-DPPO(1) catalyzed by sEH corresponds to the fourth step, the proton transfer from the water molecule to His523 with concomitant attacking of the oxygen water molecule to Asp333 and breaking the bond established with the epoxide (TS4(B2 & B3) in Scheme 4), in the case of the reaction catalyzed by Ser105Asp CALB, the barrier of this step is dramatically reduced (see the corresponding PES in Figure S3 of the Supplementary Materials) and the rate-limiting step is then associated with the nucleophilic attack of the aspartate residue on the carbon atom of the epoxide ring, the breaking of the epoxide three-membered ring, and the stabilization of the negative charge developed in the oxygen atom by the oxyanion hole.

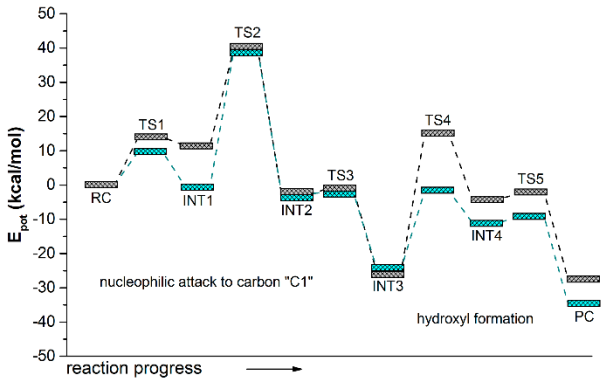


Figure 5. Potential energy profiles computed at the AM1 level for t-DPPO(1) hydrolysis catalyzed by sEH (**grey line**) and by Ser105Asp CALB (**green line**). The reaction corresponds to the catalytic attack on the C1 carbon of epoxide.

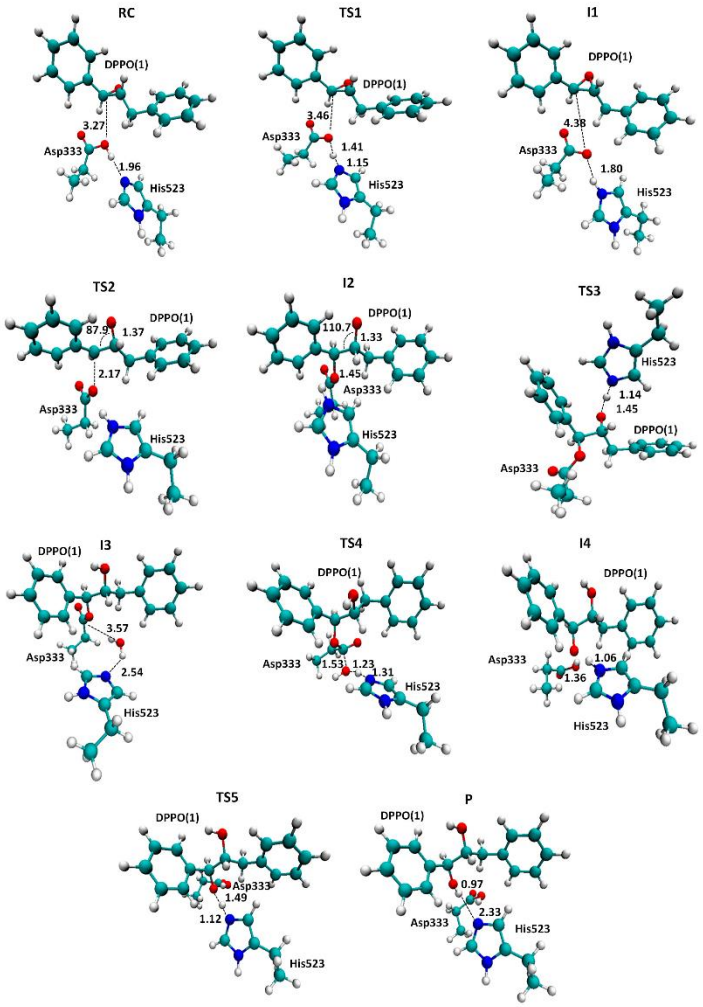


Figure 6. Localized structures for hydrolysis reaction of t-DPPO(1) catalyzed by sEH. Key inter-atomic distances are reported in Å. Tyr381 and Tyr465, forming the oxyanion hole, are not displayed for the purposes of clarity.

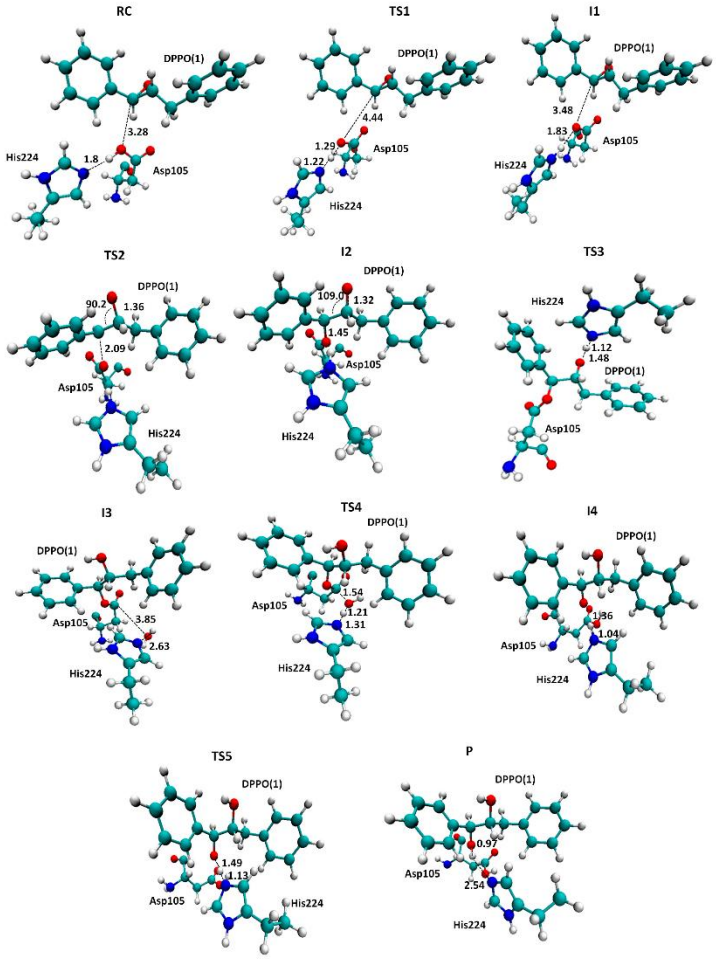


Figure 7. Localized structures for hydrolysis reaction of t-DPPO(1) catalyzed by Ser105Asp CALB. Key inter-atomic distances are reported in Å. Thr40 and Gln106, forming the oxyanion hole, are not displayed for the purposes of clarity.

3. Computational Methods

The theoretical study of possible t-DPPO epoxide hydrolysis mechanisms has been performed within two different active sites models, as presented in Scheme 1. Models were prepared based on available

Protein Data Bank crystal structures of sEH (PDB ID 1EK1) [28] and CALB (PDB ID 1TCA) [42] enzymes. Mutation of Ser105 to Asp in the CALB active site was achieved by using Discovery Studio 3.5 [43]. The cluster models, depicted in Scheme 1, have been selected to mimic the conserved catalytic triad and oxyanion hole found in the active site of these two α/β hydrolases, which have shown significant activity for epoxide hydrolysis reactions. Thus, in the model of the active site of sEH, the substrate is surrounded by Asp333, His523, Asp495, Tyr381, and Tyr465. All residues are saturated in Cα positions. The PESs for the different mechanism have been obtained at the AM1 semi-empirical level [44]. The highly parametrized M06-2X [45,46] hybrid functional was selected in order to improve the limitation of the AM1 method. The 6-31+G** basis set was used for the DFT calculations. After localizing the stationary points, frequency calculations were carried out to verify that the structures represent true minima or first-order saddle points on the gas phase PESs. Once first-order saddle points were located and characterized, the Intrinsic Reaction Coordinate (IRC) path was traced down from the saddle points to the corresponding minima using the full gradient vector. The global r.m.s. residual gradient in the optimized structures was always less than 0.04 $kcal \cdot mol^{-1} \cdot Å^{-1}$. It is important to note that no constraints were applied to any of the geometry optimizations. Although allowing more reliable energetics, this implies that possible artifacts, such as odd interaction complexes, can be obtained. Thus, proper orientation of the different structures in the starting point structures is a crucial step in the computational protocol. Also, keeping in mind that the reaction under study is a multi-step process, IRC calculations traced forward from a TS structure do not necessarily converge in the end of the backwards path traced from the following IRC. In this sense, efforts have been made to get a converged result; otherwise, the minimum energy structure, belonging to the reaction path, was selected. All AM1 calculations were performed with MOPAC2007 [47], while DFT calculations were with Gaussian 09 [48].

4. Conclusions

The primary reaction of the sEH, the hydrolysis of epoxides, has been studied by means of the AM1 semi-empirical method and the M06-2X hybrid functional, using a cluster model to mimic its active site. Using this reaction as a template, the secondary activity of wild-type CALB as an epoxidase has been studied and the comparison with the reaction on sEH has suggested a mutation of one residue of the active site that could improve the catalytic activity of CALB. Indeed, our results carried out with t-DPPO as a substrate suggest that the multi-step mechanism that Ser105Asp employs to hydrolyze t-DPPO is similar to the one used by sEH. Our predicted rate limiting step of sEH is in agreement with the pre-steady-state kinetic analysis of epoxide hydration catalyzed by mEH performed by Tseng *et al.* [49], who concluded that the rate of hydrolysis of the hydroxyl alkyl-enzyme intermediate was far slower than the rate of its formation. When the reaction is studied in the active site of the Ser105Asp CALB, our calculations show that the barrier of this step is dramatically reduced and the step associated to the nucleophilic attack of the aspartate residue to the carbon atom of the epoxide ring becomes the rate-limiting step. The barrier of this step is almost equivalent in both sEH and Ser105Asp CALB. Finally, any attempt to perform the nucleophilic attack by a possible water molecule in the active site of the sEH results in higher energy barriers, thus discarding this alternative mechanism. At this point, we must keep in mind that the values of the relative energies of the stationary points can be shifted when including the

effect of the full protein. Nevertheless, since a comparative study has been carried out, in general the consequences should not be too dramatic.

From the computational point of view, our results, obtained at the AM1 level, follow the same trend as the ones recomputed at a higher level of theory, M06-2X with the 6-31+G** basis set. These latter values are, in fact, quite close to the barriers that can be deduced from experimentally measured rate constant in related enzymatic processes. Nevertheless, it appears that the stereo-selectivity of the enzyme, as a result of the nucleophilic attack taking place on one or the other carbon atom of the epoxide, cannot be predicted from simulations based on reduced models. Thus, further studies should be carried out in the future with more complex and realistic models such as the ones based on the use of hybrid QM/MM models. Information derived from the present study will be of great help in setting up the model, in determining the size and level of theory to be employed in the QM region, and in the analysis of the role of the key residues of the active site.

Supplementary Materials

Supplementary materials can be accessed at: http://www.mdpi.com/1420-3049/20/10/17789/s1.

Acknowledgments

This work was supported by the Spanish Ministerio de Economía y Competitividad for project CTQ2012-36253-C03, Universitat Jaume I (project P1•1B2014-26), Generalitat Valenciana (PROMETEOII/2014/022 and ACOMP/2014/277 projects), the Polish National Center for Science (NCN) (grant 2011/02/A/ ST4/00246, 2012–2017), the Polish Ministry of Science and Higher Education ("Iuventus Plus" program project no. 0478/IP3/2015/73, 2015-2016), and the USA's National Institutes of Health (ref. NIH R01 GM065368). The authors acknowledge computational resources from the Servei d'Informàtica of Universitat Jaume I and from the Technical University of Lodz.

Author Contributions

I.B.—calculation for sEH model and figures preparation; J.R.G.—calculations for CALB model; K.Ś.—design of research, setting up of molecular models, results analysis, figures preparation, and writing manuscript; V.M.—design of research, results analysis, and writing manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Devries, E.J.; Janssen, D.B. Biocatalytic conversion of epoxides. *Curr. Opin. Biotechnol.* **2003**, *14*, 414–420.
2. Parker, R.E.; Isaacs, N.S. Mechanism of epoxide reactions. *Chem. Rev.* **1959**, *264*, 9310–9313.
3. Jacobsen, E.N. Asymmetric catalysis of epoxide ring-opening reactions. *Acc. Chem. Res.* **2000**, *33*, 421–431.

4. Archelas, A.; Furstoss, R. Synthesis of enantiopure epoxides through biocatalytic approaches. *Annu. Rev. Microbiol.* **1997**, *51*, 491–525.
5. Archelas, A.; Furstoss, R. Synthetic applications of epoxide hydrolases. *Curr. Opin. Chem. Biol.* **2001**, *5*, 112–119.
6. Wixtrom, R.N.; Hammock, B.D. Membrane-bound and soluble-fraction epoxide hydrolases: Methodological aspects. In *Biochemical Pharmacology and Toxicology: Methodological Aspects of Drug Metabolizing Enzymes*; Akim, D.Z., Vessey, D.A., Eds.; Wiley: New York, NY, USA, 1985; Volume 1, pp. 1–93.
7. Moghaddam, M.F.; Grant, D.F.; Cheek, J.M.; Greene, J.F.; Williamson, K.C.; Hammock, B.D. Bioactivation of leukotoxins to their toxic diols by epoxide hydrolase. *Nat. Med.* **1997**, *3*, 562–566.
8. Yu, Z.; Xu, F.; Huse, L.M.; Morisseau, C.; Draper, A.J.; Newman, J.W.; Parker, C.; Graham, L.; Engler, M.M.; Hammock, B.D.; *et al.* Soluble epoxide hydrolase regulates hydrolysis of vasoactiveepoxyeicosatrienoic acids. *Circ. Res.* **2000**, *87*, 992–998.
9. Sinal, C.J.; Miyata, M.; Tohkin, M.; Nagata, K.; Bend, J.R.; Gonzalez, F.J. Targeted disruption of soluble epoxide hydrolase reveals a role in blood pressure regulation. *J. Biol. Chem.* **2000**, *275*, 40504–40510.
10. Campbell, W.B. New role for epoxyeicosatrienoic acids as anti-inflammatory mediators. *Trends Pharmacol. Sci.* **2000**, *21*, 125–127.
11. Davis, B.B.; Thompson, D.A.; Howard, L.L.; Morisseau, C.; Hammock, B.D.; Weiss, R.H. Inhibitors of soluble epoxide hydrolase attenuate vascular smooth muscle cell proliferation. *Proc. Natl. Acad. Sci USA* **2002**, *99*, 2222–2227.
12. Imig, J.D.; Zhao, X.; Capdevila, J.H.; Morisseau, C.; Hammock, B.D. Soluble epoxide hydrolase inhibition lowers arterial blood pressure in angiotensin II hypertension. *Hypertension* **2002**, *39*, 690–694.
13. Yu, Z.; Davis, B.B.; Morisseau, C.; Hammock, B.D.; Olson, J.L.; Kroetz, D.L.; Weiss, R.H. Vascular localization of soluble epoxide hydrolase in humankidney. *Am. J. Physiol. Renal Physiol.* **2003**, *286*, F720–F726.
14. Zhao, X.; Yamamoto, T.; Newman, J.W.; Kim, I.H.; Watanabe, T.; Hammock, B.D.; Stewart, J.; Pollock, J.S.; Pollock, D.M.; Imig, J.D. Soluble epoxide hydrolase inhibition protects the kidney from hypertension-induced damage. *J. Am. Soc. Nephrol.* **2004**, *15*, 1244–1253.
15. Imig, J.D.; Hammock, B.D. Soluble epoxide hydrolase as a therapeutic target for cardiovascular diseases. *Nat. Rev. Drug Discov.* **2009**, *8*, 794–805.
16. Wang, Z.H.; Davis, B.B.; Jiang, D.Q.; Zhao, T.T.; Xu, D.Y. Soluble epoxide hydrolase inhibitors and cardiovascular diseases. *Curr. Vasc. Pharmacol.* **2013**, *11*, 105–111.
17. Holmquist, M. α/β-Hydrolase fold enzymes: Structures, functions and mechanisms. *Curr. Protein Pept. Sci.* **2000**, *1*, 209–235.
18. Beetham, J.K.; Grant, D.; Arand, M.; Garbarino, J.; Kiyosue, T.; Pinot, F.; Oesch, F.; Belknap, W.R.; Shinozaki, K.; Hammock, B.D. Gene evolution of epoxide hydrolases and recommended nomenclature. *DNA Cell Biol.* **1995**, *14*, 61–71.
19. Yamada, T.; Morisseau, C.; Maxwell, J.E.; Argiriadi, M.A.; Christianson, D.W.; Hammock, B.D. Biochemical evidence for the involvement of tyrosine in epoxide activation during the catalytic cycle of epoxide hydrolase. *J. Biol. Chem.* **2000**, *275*, 23082–23088.

20. Busto, E.; Gotor-Fernández, V.; Gotor, V. Hydrolases: Catalytically promiscuous enzymes for non-conventional reactions in organic synthesis. *Chem. Soc. Rev.* **2010**, *39*, 4504–4523.
21. Svedendahl, M.; Jovanovic, B.; Fransson, L.; Berglund, P. Suppressed native hydrolytic activity of a lipase to reveal promiscuous Michael addition activity in water. *ChemCatChem* **2009**, *1*, 252–258.
22. Świderek, K.; Martí, S.; Moliner, V. Theoretical study of primary reaction of Pseudozymaantarctica lipase B as the starting point to understand its promiscuity. *ACS Catal.* **2014**, *4*, 426–434.
23. Ollis, D.L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S.M.; Harel, M.; Remington, S.J.; Silman, I.; Schrag, J.; *et al.* The α/β hydrolase fold. *Protein Eng.* **1992**, *5*, 197–211.
24. Świderek, K.; Pabis, A.; Moliner, V. A theoretical study of carbon-carbon bond formation by a Michael-type addition. *Org. Biomol. Chem.* **2012**, *10*, 5598–5605.
25. Chen, B.; Hu, J.; Miller, E.; Xie, W.; Cai, M.; Gross, R. Candida antarctica Lipase B chemically immobilized on epoxy-activated micro- and nanobeads: Catalysts for Polyester Synthesis. *Biomacromolecules* **2008**, *9*, 463–471.
26. Svedendahl, M.; Carlqvist, P.; Branneby, C.; Allnér, O.; Frise, A.; Hult, K.; Berglund, P.; Brinck, T. Direct epoxidation in *Candida antarctica* lipase B studied by experiment and theory. *ChemBioChem* **2008**, *9*, 2443–2451.
27. Lonsdale, R.; Hoyle, S.; Grey, D.T.; Ridder, L.; Mulholland, A.J. Determinants of reactivity and selectivity in soluble epoxide hydrolase from quantum mechanics/molecular Mechanics modeling. *Biochemistry* **2012**, *51*, 1774–1786.
28. Argiriadi, M.A.; Morisseau, C.; Goodrow, M.H.; Dowdy, D.L.; Hammock, B.D.; Christianson, D.W. Binding of alkylurea inhibitors to epoxide hydrolase implicates active site tyrosines in substrate activation. *J. Biol. Chem.* **2000**, *275*, 15265–15270.
29. Schiøtt, B.; Bruice, T.C. Reaction mechanism of soluble epoxide hydrolase: Insights from molecular dynamics simulations. *J. Am. Chem. Soc.* **2002**, *124*, 14558–14570.
30. Amrein, B.A.; Bauer, P.; Duarte, F.; Carlsson, Å.J.; Naworyta, A.; Mowbray, S.L.; Widersten, M.; Kamerlin, S.C.L. Expanding the catalytic triad in epoxide hydrolases and related enzymes. *ACS Catal.* **2015**, *5*, 5702–5713.
31. Borhan, B.; Jones, A.D.; Pinot, F.; Grant, D.F.; Kurth, M.J.; Hammock, B.D. Mechanism of soluble epoxide hydrolase: Formation of an α-hydroxy ester-enzyme intermediate through Asp-333. *J. Biol. Chem.* **1995**, *270*, 26923–26930.
32. Li, H.; Robertson, A.D.; Jensen, J.H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **2005**, *61*, 704–721.
33. Bas, D.C.; Rogers, D.M.; Jensen, J.H. Very fast prediction and rationalization of pK_a values for protein-ligand complexes. *Proteins* **2008**, *73*, 765–783.
34. Olsson, M.H.M.; Søndergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
35. Søndergaard, C.R.; Olsson, M.H.M.; Rostkowski, M.; Jensen, J.H.J. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pka values. *Chem. Theory Comput.* **2011**, *7*, 2284–2295.

36. Hopmann, K.H.; Himo, F. Theoretical study of the full reaction mechanism of human soluble epoxide hydrolase. *Chem. Eur. J.* **2006**, *12*, 6898–6909.
37. Hopmann, K.H.; Himo, F. Insights into the reaction mechanism of soluble epoxide hydrolase from theoretical active site mutants. *J. Phys. Chem. B* **2006**, *110*, 21299–21310.
38. Armstrong, R.N. Enzyme-catalyzed detoxification reactions: Mechanisms and stereochemistry. *Crit. Rev. Biochem. Mol. Biol.* **1987**, *22*, 39–88.
39. Pinot, F.; Grant, D.F.; Beetham, J.K.; Parker, A.G.; Borhan, B.; Landt, S.; Jones, A.D.; Hammock, B.D. Molecular and biochemical evidence for the involvement of the Asp333-His523 pair in the catalytic mechanism of soluble epoxide hydrolase. *J. Biol. Chem.* **1995**, *270*, 7968–7974.
40. Morisseau, C.; Du, G.; Newman, J.W.; Hammock, B.D. Mechanism of mammalian soluble epoxide hydrolase inhibition by chalcone oxide derivatives. *Arch. Biochem. Biophys.* **1998**, *356*, 214–228.
41. Elfström, L.T.; Widersten, M. Implications for an ionized alkyl-enzyme intermediate during StEH1-catalyzed *trans*-stilbene oxide hydrolysis. *Biochemistry* **2006**, *45*, 205–212.
42. Uppenberg, J.; Hansen, M.T.; Patkar, S.; Jones, T.A. The sequence, crystal structure determination and refinement of two crystal forms of lipase B from *Candida antarctica*. *Structure* **1994**, *2*, 293–308.
43. *Discovery Studio Modeling Environment*, Release 3.5; Dassault Systèmes BIOVIA: San Diego, CA, USA, 2015.
44. Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. The development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
45. Zhao, Y.; Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
46. Zhao, Y.; Truhlar, D.G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.
47. Stewart, J.J.P. *MOPAC2007*; Computational Chemistry: Colorado Springs, CO, USA, 2007.
48. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; *et al*. *GAUSSIAN 09*, Revision A.1; Gaussian, Inc.: Wallingford, CT, USA, 2009.
49. Tzeng, H.F.; Laughlin, L.T.; Lin, S.; Armstrong, R.N. The catalytic mechanism of microsomal epoxide hydrolase involves reversible formation and rate limiting hydrolysis of the alkyl-enzyme intermediate. *J. Am. Chem. Soc.* **1996**, *118*, 9436–9437.

Sample Availability: Not available.

Supplementary Material

Table S1. pKa values of key residues in the active site of sEH and Ser105Asp CALB as computed with PROPKA program, compared with their standard values in solution.

	Residue	pKa in Enzyme	Standard pKa
sEH	Asp 333	10.31	3.80
	Asp 495	9.08	3.80
	His 523	1.23	6.50
	Tyr 465	18.12	10.00
	Tyr 381	14.82	10.00
Ser105Asp CALB	Asp 105	11.41	3.80
	His 224	5.19	6.50
	Asp 187	2.80	3.80

Table S2. Important distances for nucleophilic attack of water to C1 of DPPO(1) catalyzed by sEH. Values are given in Å.

Distances	RC	TS1	I1	TS2	PC
$H\delta^{His523}$–$N\delta^{His523}$	1.01	1.01	1.04(1.03)	1.02	1.01
$H\delta^{His523}$–O^{Asp495}	1.95	1.90	1.78(1.82)	1.85	1.94
O^{epox}–$C1^{epox}$	1.45	2.17	2.41(2.41)	2.41	2.42
O^{epox}–$C2^{epox}$	1.45	1.34	1.33(1.35)	1.37	1.43
O^{epox}–H^{Tyr381}	2.13	1.94	1.88(1.89)	1.92	2.10
O^{epox}–H^{Tyr465}	2.15	1.99	1.94(1.92)	1.97	2.23

Table S3. Important distances for nucleophilic attack of Aspartate to C1of DPPO(1) catalyzed by sEH and Ser105CALB. Values are given in Å.

Distances	RC		TS1		I1		TS2		I2	
	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB
O^{Asp333}–H^{Asp333}	0.98	0.99	1.41	1.29	1.80(1.83)	1.83 (2.23)	1.88	1.90	2.45	2.30
H^{Asp333}–$N\varepsilon^{His523}$	1.96	1.87	1.15	1.22	1.03(1.03)	1.02 (1.02)	1.02	1.01	1.00	1.00
$H\delta^{His523}$–$N\delta^{His523}$	1.01	1.02	1.03	1.03	1.03(1.04)	1.04 (1.03)	1.04	1.04	1.04	1.04
$H\delta^{His523}$–O^{Asp495}	1.93	2.02	1.85	1.96	1.80(1.79)	1.98(1.95)	1.78	1.92	1.77	1.93
O^{Asp333}–$C1^{epox}$	3.27	3.28	3.46	4.44	4.38(3.26)	3.48 (3.94)	2.17	2.09	1.45	1.45
O^{Asp333}–$C2^{epox}$	3.58	4.31	3.73	5.60	4.10(3.56)	4.41 (4.06)	2.93	2.88	2.43	2.46
O^{epox}–$C1^{epox}$	1.45	1.44	1.45	1.45	1.45(1.45)	1.44 (1.44)	1.99	2.03	2.40	2.38
O^{epox}–$C2^{epox}$	1.44	1.44	1.44	1.44	1.45(1.44)	1.44 (1.44)	1.37	1.36	1.33	1.32
O^{epox}–H^{Tyr185}	2.13	2.25	2.12	2.24	2.14(2.12)	2.24(2.20)	1.98	1.99	1.90	1.91
O^{epox}–H^{Tyr185}	2.14	2.20	2.14	2.21	2.17(2.13)	2.19(2.36)	1.97	2.20	1.92	2.19

Table S4. Important distances for nucleophilic attack of Aspartate to C2 of DPPO(1) catalyzed by sEH and Ser105CALB.Values are given in Å.

Distances	RC		TS1		I1		TS2		I2	
	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB
O^{Asp333}–H^{Asp333}	0.98	0.99	1.41	1.29	1.80(2.05)	1.83 (1.86)	2.10	1.95	2.44	2.30
H^{Asp333}–$N\varepsilon^{His523}$	1.96	1.87	1.15	1.22	1.03(1.01)	1.02 (1.02)	1.01	1.01	1.00	1.00
$H\delta^{His523}$–$N\delta^{His523}$	1.01	1.02	1.03	1.03	1.03(1.03)	1.04 (1.04)	1.04	1.04	1.04	1.04
$H\delta^{His523}$ O^{Asp495}	1.93	2.02	1.85	1.96	1.80()	1.98(1.92)	1.80	1.93	1.77	1.95
O^{Asp333}–$C1^{epox}$	3.27	3.28	3.46	4.44	4.38(3.98)	3.48 (4.82)	3.02	2.93	2.54	2.55
O^{Asp333}–$C2^{epox}$	3.58	4.31	3.76	5.60	4.10(3.15)	4.41 (4.07)	2.11	1.98	1.44	1.45
O^{epox} $C1^{epox}$	1.45	1.44	1.45	1.45	1.45(1.45)	1.44 (1.44)	1.38	1.35	1.33	1.32
O^{epox}–$C2^{epox}$	1.44	1.44	1.44	1.44	1.45(1.45)	1.44 (1.44)	1.96	2.05	2.37	2.42
O^{epox}–H^{Tyr185}	2.13	2.25	2.12	2.24	2.14(2.15)	2.24(4.50)	1.98	4.17	1.90	1.93
O^{epox}–H^{Tyr185}	2.14	2.20	2.14	2.21	2.17(2.13)	2.19(6.35)	1.99	5.69	1.86	2.26

Table S5. Important distances for nucleophilic attack of Aspartate to C1 of DPPO(2) catalyzed by sEH and Ser105CALB.Values are given in Å.

Distances	RC		TS1		I1		TS2		I2	
	Wild Type sEH	Scr105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB
O^{Asp333}–H^{Asp333}	0.98	0.99	1.43	1.26	1.79(1.84)	1.86(1.87)	1.90	1.93	2.07	2.12
H^{Asp333}–$N\varepsilon^{His523}$	2.12	1.87	1.14	1.25	1.03(1.02)	1.02(1.02)	1.02	1.01	1.00	1.00
$H\delta^{His523}$–$N\delta^{His522}$	1.01	1.02	1.03	1.02	1.04(1.04)	1.04(1.04)	1.04	1.04	1.05	1.04
$H\delta^{His522}$–O^{Asp195}	1.91	2.01	1.85	1.95	1.80(1.79)	1.90(2.01)	1.77	1.91	1.74	1.90
O^{Asp333}–$C1^{spvs}$	3.63	4.56	3.96	4.83	3.47(3.70)	4.62(3.78)	2.09	1.99	1.45	1.46
O^{Asp333}–$C2^{spvs}$	3.77	3.48	3.81	3.76	3.65(3.25)	3.62(3.85)	2.97	2.95	2.58	2.50
O^{spvs}–$C1^{spvs}$	1.44	1.44	1.45	1.44	1.44(1.45)	1.44(1.44)	1.96	2.08	2.37	2.43
O^{spvs}–$C2^{spvs}$	1.45	1.44	1.45	1.44	1.45(1.45)	1.44(1.44)	1.38	1.34	1.33	1.32
O^{-pvs}–H^{Tyr161}	2.15	2.46	2.15	2.50	2.13(2.15)	2.65(5.60)	2.00	3.99	1.93	1.91
O^{-pvs}–H^{Tyr165}	2.13	3.69	2.12	3.56	2.12(2.13)	3.81(6.43)	1.98	5.57	1.89	2.18

Table S6. Important distances for nucleophilic attack of Aspartate to C2 of DPPO(2) catalyzed by sEH and Ser105CALB.Values are given in Å.

Distances	RC		TS1		I1		TS2		I2	
	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB
O^{Asp333}–H^{Asp333}	0.98	0.99	1.43	1.26	1.79(1.77)	1.86(2.27)	1.88	2.04	2.16	2.21
H^{Asp333}–$N\varepsilon^{His553}$	2.12	1.87	1.14	1.25	1.03(1.03)	1.02(1.02)	1.02	1.01	1.01	1.00
$H\delta^{His523}$–$N\delta^{His522}$	1.01	1.02	1.03	1.02	1.04(1.03)	1.04(1.03)	1.04	1.04	1.04	1.04
$H\delta^{His522}$–O^{Asp195}	1.91	2.01	1.85	1.95	1.80(1.80)	1.90(1.78)	1.78	1.94	1.76	1.76
O^{Asp333}–$C1^{spvs}$	3.63	4.56	3.96	4.83	3.47(3.61)	4.62(4.45)	2.94	2.90	2.43	2.51
O^{Asp333}–$C2^{spvs}$	3.77	3.48	3.81	3.76	3.65(2.92)	3.62(3.46)	2.11	2.00	1.45	1.44
O^{spvs}–$C1^{spvs}$	1.44	1.44	1.45	1.44	1.44(1.45)	1.44(1.44)	1.38	1.36	1.33	1.32
O^{spvs}–$C2^{spvs}$	1.45	1.44	1.45	1.44	1.45(1.45)	1.44(1.45)	1.96	2.01	2.39	2.36
O^{-pvs}–H^{Tyr161}	2.15	2.46	2.15	2.50	2.13(2.15)	2.65(2.14)	2.00	2.09	1.94	1.94
O^{spvs}–H^{Tyr165}	2.13	3.69	2.12	3.56	2.12(2.15)	3.81(2.26)	1.98	2.15	1.89	2.13

S4

Table S7. Important distances for hydroxyl formation for DPPO(1) catalyzed by sEH. Values are given in Å.

Distances	I2		TS3(B1)		I3		TS4(B1)		P	
	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB	Wild Type sEH	Ser105Asp CALB
O^{Asp332}–H^{Asp332}	2.22	2.27	2.18	2.66	2.09(4.86)	4.66 (4.66)	4.64	4.56	3.60	4.08
H^{Asp332}–$N\delta^{His523}$	1.00	1.00	1.00	1.00	1.00(1.04)	1.04 (1.05)	1.11	1.10	2.59	2.60
$H\delta^{His523}$–$N\delta^{His523}$	1.04	1.04	1.04	1.05	1.04(1.03)	1.04 (1.03)	1.03	1.03	1.01	1.01
$H\delta^{His523}$–O^{Asp395}	1.76	1.76	1.75	1.73	1.76(1.81)	1.79(1.94)	1.83	1.96	1.98	2.04
O^{Asp335}–$C1^{epox}$	1.44	1.45	1.43	1.45	1.43(1.43)	1.43 (1.42)	1.43	1.43	1.43	1.42
O^{Asp335}–$C2^{epox}$	2.48	2.48	2.50	2.44	2.52(2.52)	2.84 (2.49)	2.48	2.48	2.45	2.44
O^{epox}–$C1^{epox}$	2.37	2.37	2.38	2.40	2.38(2.38)	2.43 (2.40)	2.38	2.40	2.41	2.41
O^{epox}–$C2^{epox}$	1.33	1.32	1.33	1.32	1.33(1.36)	1.34 (1.34)	1.37	1.36	1.43	1.42
O^{epox}–H^{Tyr381}	1.88	1.92	1.90	1.91	1.90(1.88)	1.96(1.93)	1.92	1.96	2.06	2.10
O^{epox}–H^{Tyr465}	1.91	2.16	1.92	2.20	1.90(1.97)	2.17(2.14)	2.02	2.15	2.29	2.29
O^{epox}–C^{Asn333}	3.08	3.56	1.54	1.53	1.36(1.36)	1.36 (1.36)	1.36	1.36	1.36	1.36
H^{epox}–O^{Asp250}	3.31	4.26	1.30	1.30	0.97(0.97)	0.97 (0.97)	0.97	0.97	0.97	0.97

S5

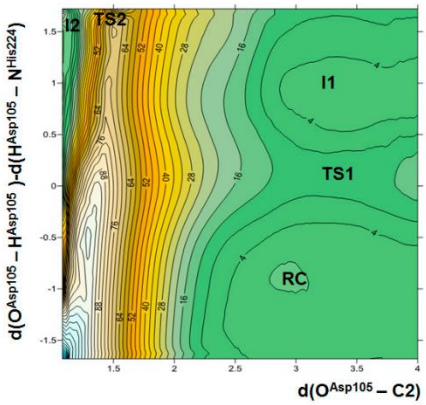


Figure S1. PES of the nucleophilic attack to C2 carbon atom of t-DPPO(1) in Ser105Asp CALB. All distances are in Å and energies in kcal·mol^{-1}.

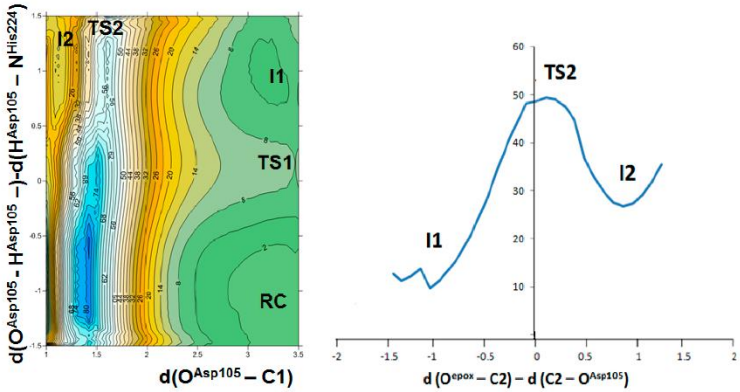


Figure S2. PES of the nucleophilic attack to C1 carbon atom of t-DPPO(2) in Ser105Asp CALB (**a**) and PES of the nucleophilic attack to C2 carbon atom (**b**). All distances are in Å and energies in kcal·mol^{-1}.

S6

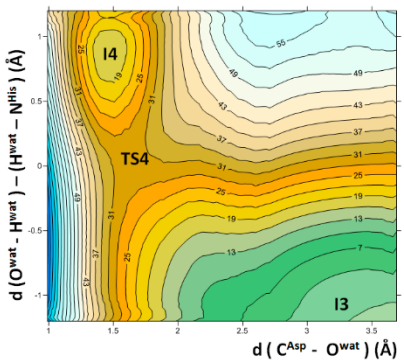


Figure S3. PES of the I3 to I4 step of t-DPPO(1) hydrolysis in Ser105Asp CALB. Distances are in Å and energies in kcal·mol^{-1}.