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

Isabel Bordes,[1] Eduardo García-Junceda,[2] Israel Sánchez-Moreno,[2] Raquel Castillo,[1,] Vicent Moliner[1,*]*

1. Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain
2. Departamento de Química Bioorgánica, Instituto de Química Orgánica General, CSIC. Juan de la Cierva 3, Madrid 28006, Spain

Corresponding authors

R. Castillo: rcastill@uji.es, tel: +34964728097

V. Moliner: moliner@uji.es, tel: +34964728084

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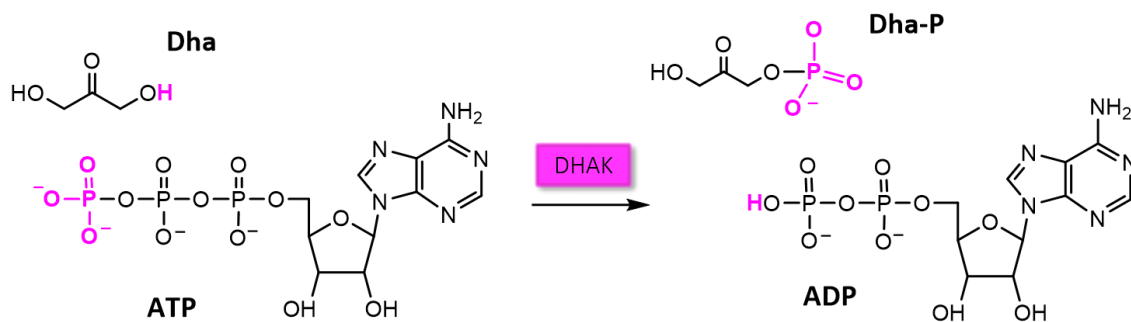
Scheme 1. Schematic representation of the reaction between ATP and Dha generating Dha-P catalysed by DHAK.

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 (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]
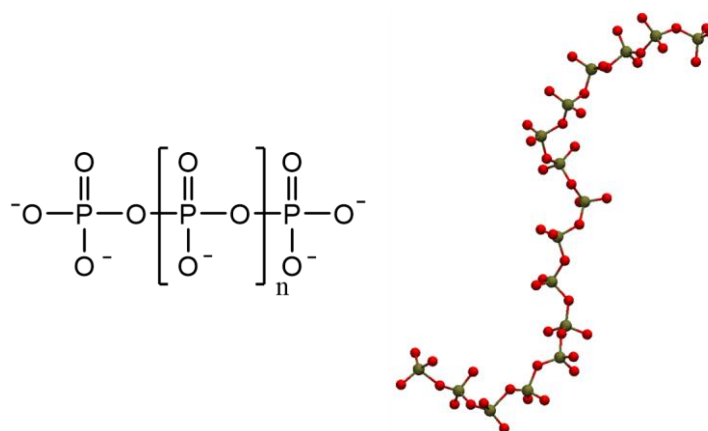

Figure 1. Chemical structure of inorganic polyphosphate polymer (left) and a ball and stick representation of a polyphosphate of n=14 (right).

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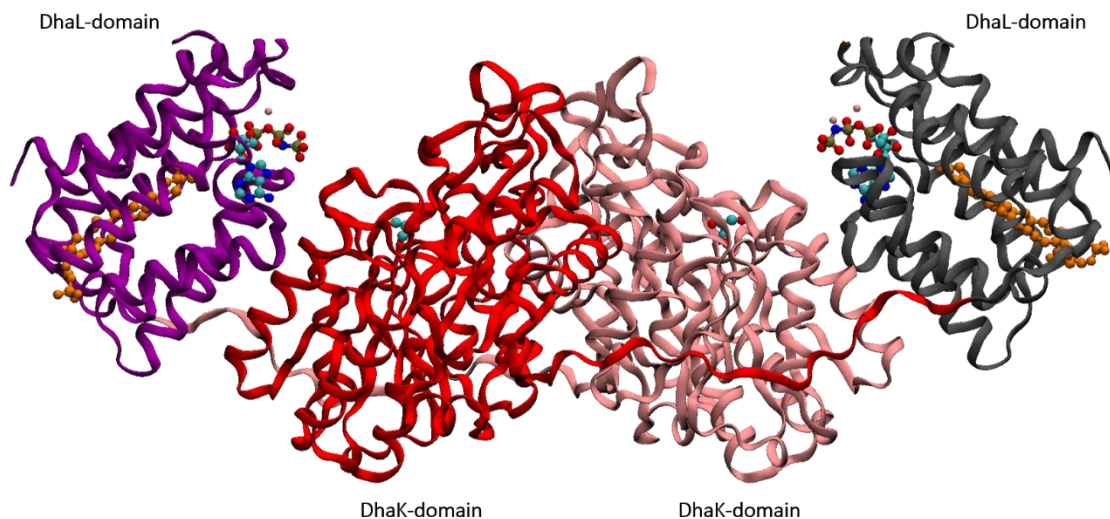
Figure 2. Cartoon representation of the kinase homodimer obtained by supeposition of PDB code crystal structures 1UN8 and 1UN9. The DhaK-domain is colored in pink and red, and the DhaL-domain is colored in purple and grey. Dha, phosphoaminophosphonic acid-adelynate ester (ANP), magnesium ions and phospholipid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

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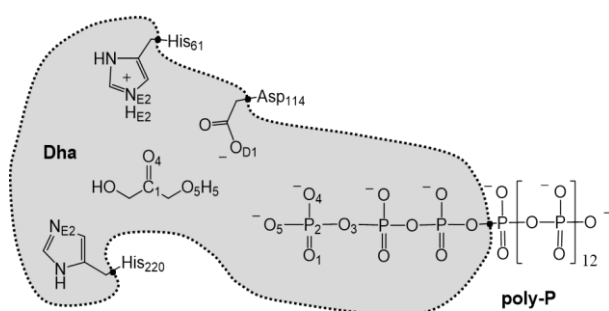
Scheme 2. Schematic representation of the active site of the DHAK. Grey region contains atoms treated quantum mechanically. Link atoms are represented as dots.

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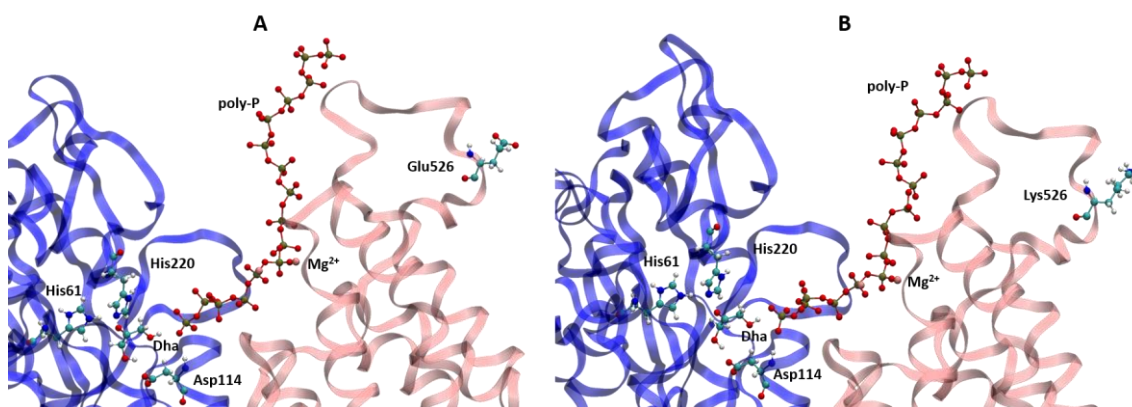
Figure 3. Active site of the wild-type (A) and the E526K mutant of *C. freundii* DHAK (B) in the reactants state conformation. Key residues (Asp114, His220, His61 and Glu/Lys526), Dha, poly-P, and Mg^{2+} ions are displayed in ball and stick representation.

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.
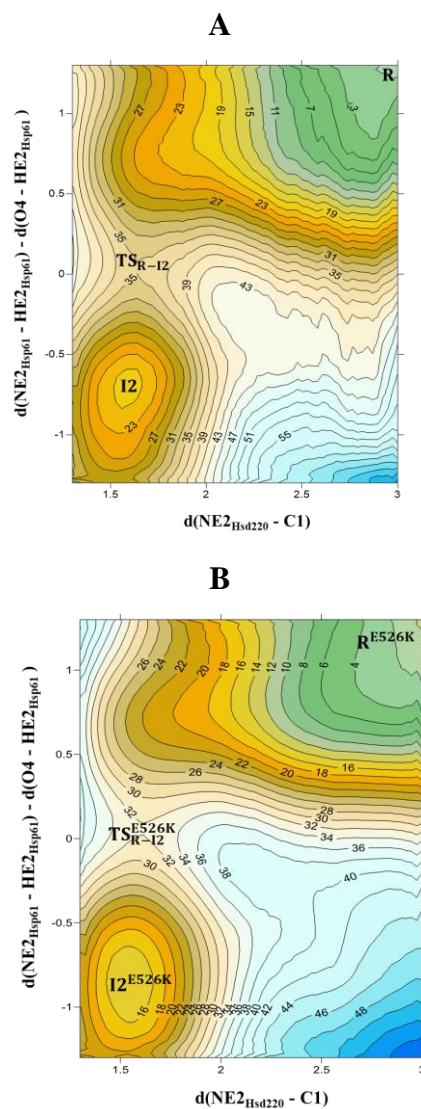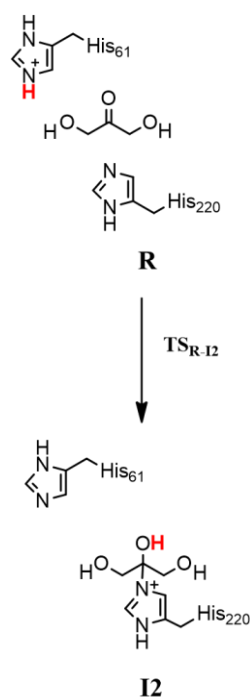
A

B

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-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, $d(O3_{poly-P} - P2_{poly-P}) - d(P2_{poly-P} - O5_{Dha})$. 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.
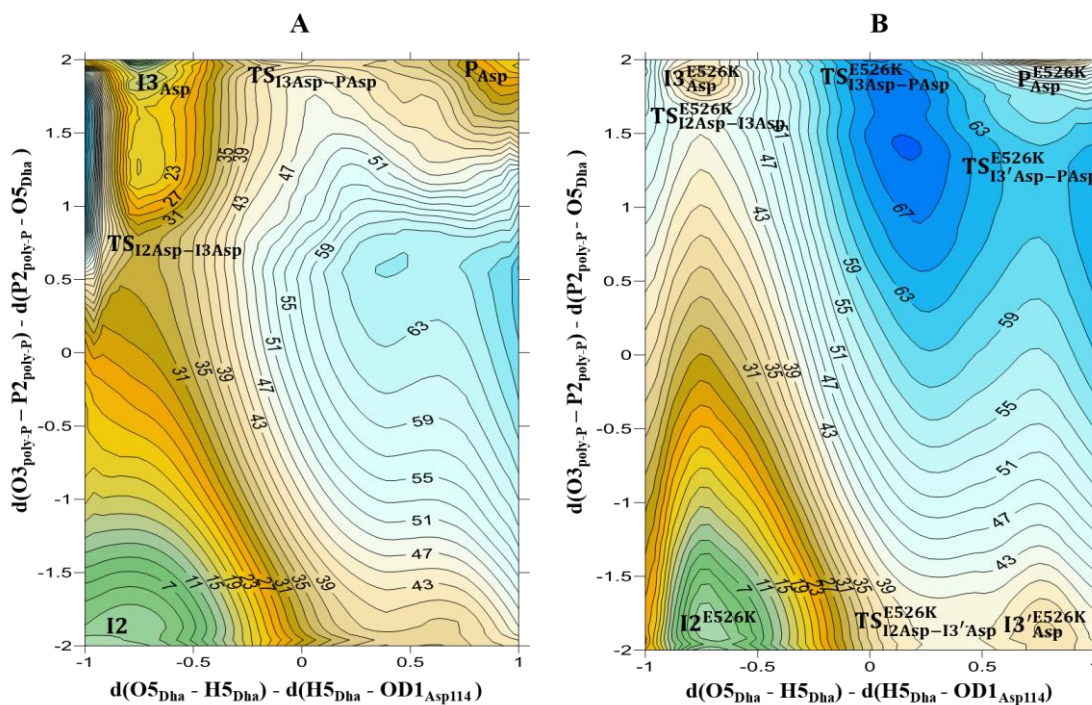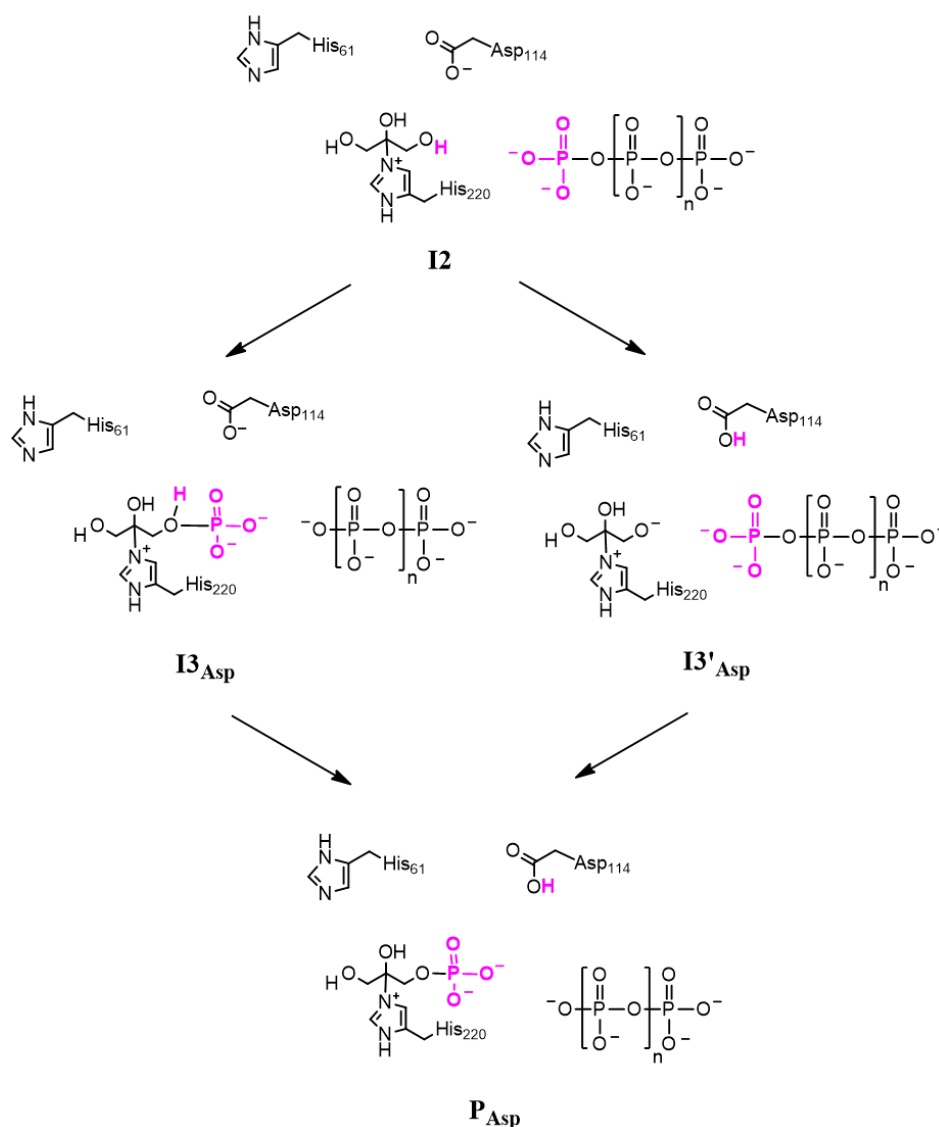

Figure 5. PM3/MM PESs of the catalytic process of the *aspartate-assisted* phosphorylation reaction mechanism from poly-P to Dha in *C. freundii* DHAK. Results are shown for the wild-type (A) and the E526K mutant (B). Distances on axis are in Å, and values of isoenergetic lines are in kcal·mol^{-1}.

I2

I3$_{Asp}$

I3'$_{Asp}$

P$_{Asp}$

Scheme 5. 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'$_{Asp}^{E526K}$. Nevertheless, despite the existence of this shallow minimum, this second path can take place in a very

14

asynchronous but concerted mechanism controlled just by the $TS^{E526K}_{I3'Asp-PAsp}$. 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 $TS^{E526K}_{I2Asp-I3Asp}$, respectively. Moreover, it must be noticed that the located intermediates, $I3_{Asp}$ and $I3^{E526K}_{Asp}$, 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 $I3^{E526K}_{Asp}$ intermediates.

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	$TS_{I2Asp-I3Asp}$	$I3_{Asp}$	$TS_{I3Asp-PAsp}$	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.
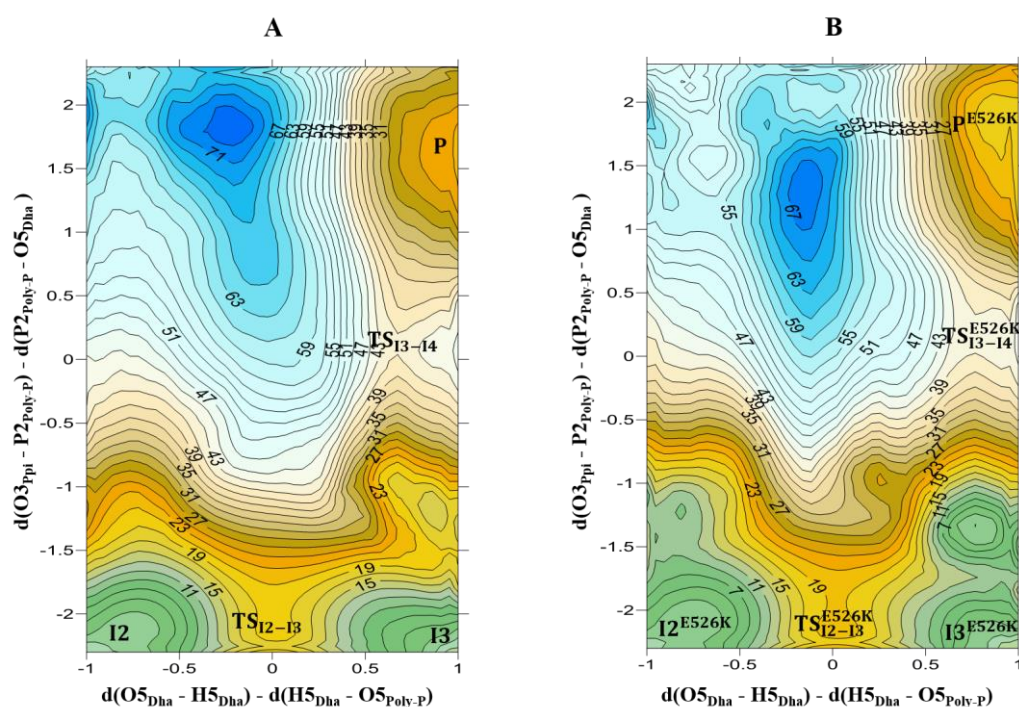

Figure 6. PM3/MM PESs of the catalytic process of the *substrate-assisted* phosphorylation reaction mechanism from poly-P to Dha in *C. freundii* DHAK. Results are shown for the wild-type (A) and the E526K mutant (B). Distances on axis are in Å, and values of isoenergetic lines are in kcal·mol^{-1}.
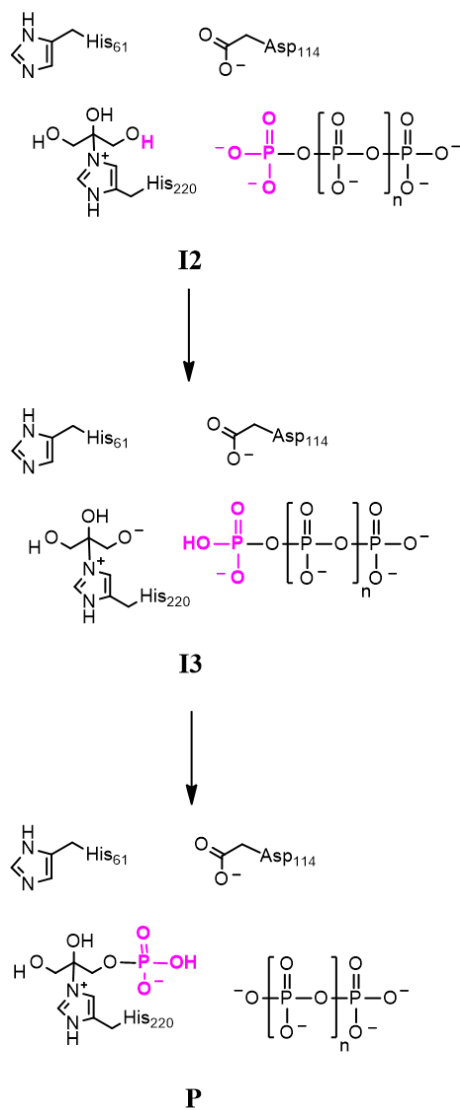
Scheme 6. Schematic representation of the *substrate-assisted mechanism* for the phosphorylation reaction from I2 to P for wild and mutant enzyme.

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$_{poly-P}$ 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$_{Dha}$-H5$_{Dha}$)	(H5$_{Dha}$-O5$_{poly-P}$)	(P2$_{poly-P}$-O3$_{poly-P}$)	(P2$_{poly-P}$-O5$_{Dha}$)	E
I2	1.04	1.62	1.71	3.77	0
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

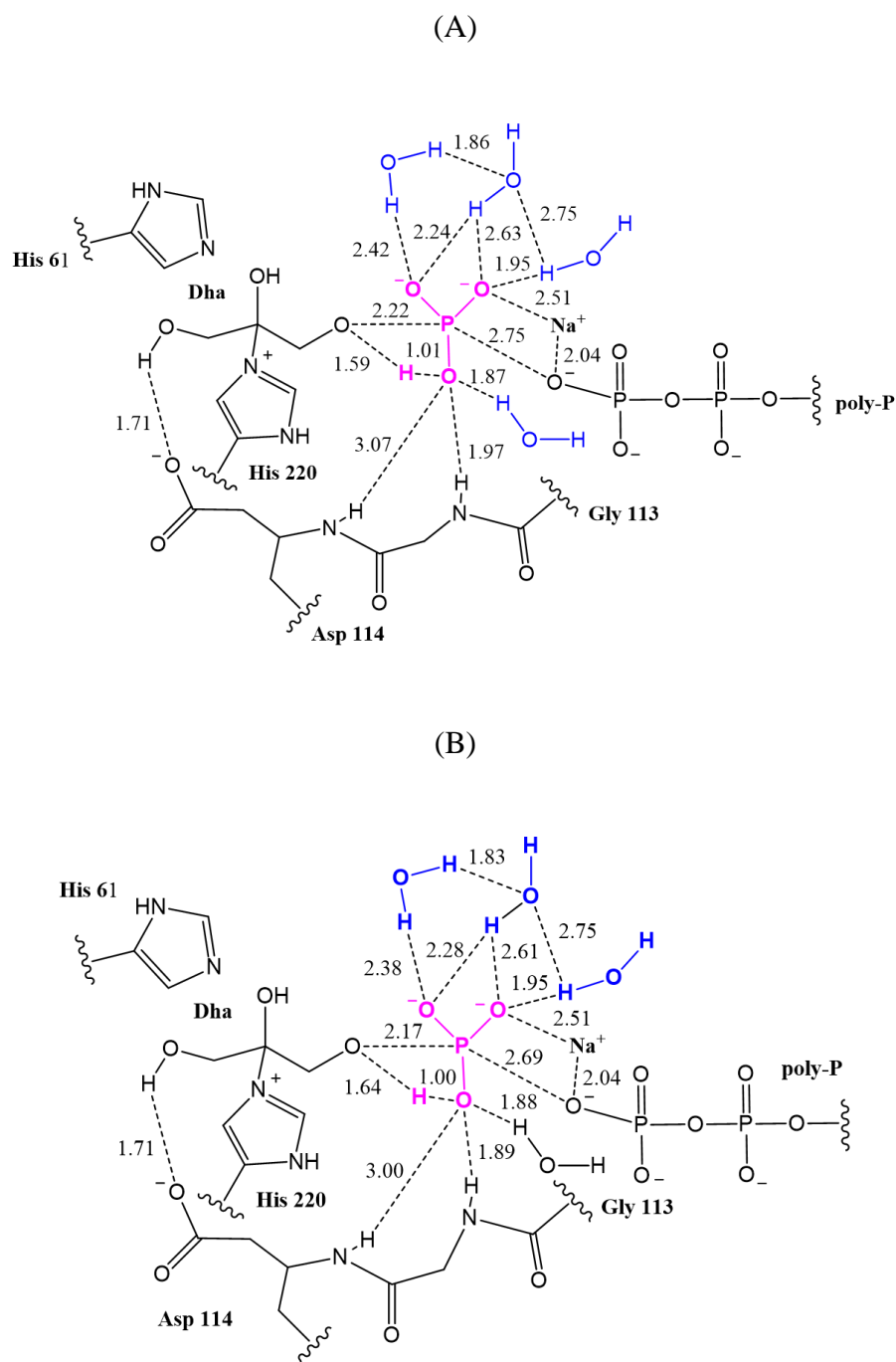|					
I2^{E526K}	1.07	1.52	1.71	3.75	0
TS$_{I2-P}^{E526K}$	1.64	1.00	2.69	2.17	71.17
P^{E526K}	1.99	0.97	3.27	1.75	64.97

(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.
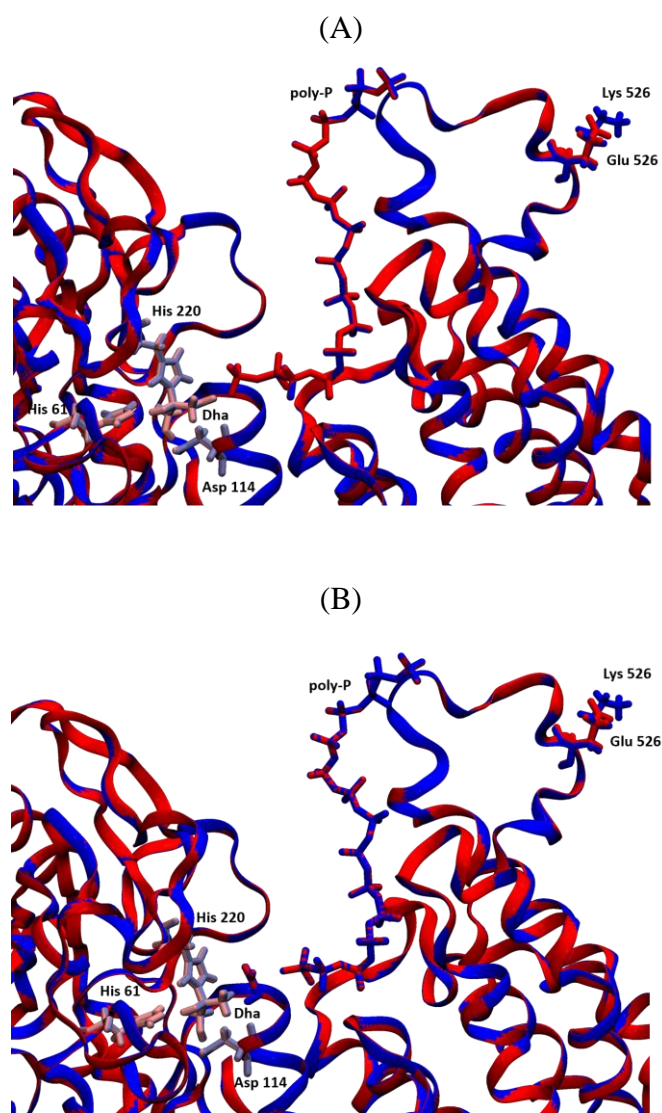
(A)

(B)

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 due to size of the systems and the fact that the study is based on explorations of PESs. In general, the analysis of energies derived from PESs for systems with a huge number of degrees of freedom has to be considered with caution. As observed in our pioneering QM/MM studies of enzyme catalyzed reactions, and in agreement with single molecule kinetic experiments,[70] the differences between the (potential energy barriers) rate constant of single molecules[71,72] can oscillate in several orders of magnitude around. A much more computationally demanding sampling of the conformational space through, for example, the use of hybrid DFT/MM molecular dynamics, can render more robust energetics. Lower energy TS conformations could be localized within this protocol thus providing much lower energy barriers. Moreover, as already observed in our previous study of the phosphoryl transfer reaction from ATP to Dha catalyzed by DhaK from Escherichia coli,[23] there must be a lack of overlapping between the B3LYP and PM3 conformational space. Then, since

the starting structures used to locate and optimize the B3LYP/MM stationary point structures were those derived from the PM3/MM calculations, they are not necessary the lowest energy ones, specially for the TSs structures. A proper sampling of the conformations on the involved states at high level of theory, i.e. the calculation of free energy barriers from statistical simulations, should be explore in future studies.

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.

SUPORTING 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

(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. *J. Phys. Chem. B* **2017**, 121, 8878–8892.
(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 Table of Contents

Herein we present 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.

Graphical Abstract