DOI: 10.1002/chem.200700044

Computational Study of the Phosphoryl Transfer Catalyzed by a Cyclin-Dependent Kinase

Marco De Vivo,[a, c] Andrea Cavalli,[a] Paolo Carloni,[b] and Maurizio Recanatini*[a]

Abstract: A cyclin-dependent kinase, Cdk2, catalyzes the transfer of the γ-phosphate from ATP to a threonine or serine residue of its polypeptide substrates. Here, we investigate aspects of the reaction mechanism of Cdk2 by gas-phase density functional calculations, classical molecular dynamics, and Car–Parrinello QM/MM simulations. We focus on the role of the conserved Asp127 and on the nature of the phosphoryl transfer reaction mechanism catalyzed by Cdk2. Our findings suggest that Asp127 is active in its deprotonated form by assisting the formation of the near-attack orientation of the substrate serine or threonine. Therefore, the residue does not act as a general base during the catalysis. The mechanism for the phosphoryl transfer is a single S_N2-like concerted step, which shows a phosphorane-like transition state geometry. Although the resulting reaction mechanism is in agreement with a previous density functional study of the same catalytic reaction mechanism (Cavalli et al., *Chem. Comm.* **2003**, 1308–1309), the reaction barrier is considerably lower when QM/MM calculations are performed, as in this study (≈42 kcal mol^{-1} QM vs. ≈24 kcal mol^{-1} QM/MM); this indicates that important roles for the catalysis are played by the protein environment and solvent waters. Because of the high amino acid sequence conservation among the whole family of cyclin-dependent kinases (CDKs), these results could be general for the CDK family.

Keywords: density functional calculations · enzymatic catalysis · molecular dynamics · reaction mechanisms · transition state analogues

Introduction

The cyclin-dependent kinase (CDK) protein family is involved in controlling the signal transduction pathways essential for the initiation, progression, and completion of the eukaryotic cell cycle.[1] Several crystallographic studies have so far elucidated the structural features of representative members of the CDK family, also providing indications about their two-step activation mechanism:[2,3] 1) the association with the requisite cyclin partner, which hence acts as regulatory subunit; 2) the phosphorylation of a specific residue, situated along the so-called activation loop (T-loop), by the CDK-activating kinase protein (CAK).[4] Once fully functionalized, CDKs transfer the γ-phosphate from ATP to a threonine or serine residue of polypeptide substrates, such as the tumor-suppressor proteins, pRb, and the related p107 and p130, and the E2F transcription factor.[5–8]

Here we focus on the enzymatic function of Cdk2, which is the most extensively studied CDK, both structurally and biochemically,[9] and a very promising target against a number of pathological conditions, including cancer.[10,11]

Cdk2's fold consists of: 1) the N-term domain (residues 1–82), mostly composed by β-sheets and only one α-helix with the typical 'PSTAIRE' sequence; 2) the hinge region (residues 83–87) connecting the N-term to the C-term; 3) the C-term domain (residues 88–298) formed principally by α-helices and by the T-loop, which contains the specific residue (Thr160) that is being phosphorylated by CAK for activation.[12,13] The catalytic site is situated at the domain interface, where the Mg^{2+} ion, the cofactor ATP, and the substrate peptide are adapted (Figure 1).

A detailed picture of the reaction mechanism and of its transition state (TS) is useful in designing TS-analogs as

[a] Dr. M. De Vivo, Dr. A. Cavalli, Prof. M. Recanatini
Department of Pharmaceutical Sciences, University of Bologna
Via Belmeloro 6, 40126 Bologna (Italy)
Fax: (+39) 051-209-9734
E-mail: maurizio.recanatini@unibo.it

[b] Prof. P. Carloni
International School for Advanced Studies SISSA/ISAS and INFM-Democritos
Via Beirut 4, 34014 Trieste (Italy)

[c] Dr. M. De Vivo
Current address: Center for Molecular Modeling (CMM)
Department of Chemistry, University of Pennsylvania
231 S. 34th Street Philadelphia, PA 19104–6323 (USA)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

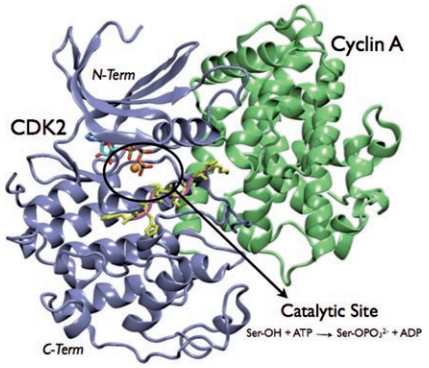


Figure 1. Cartoon representing Cdk2 (ice blue) in complex with cyclin A (green, PDB code 1QMZ).[14] ATP is represented in sticks (colored based on atom-type) and an orange sphere indicates the Mg^{2+} ion present in the active site. The peptide (HHASPRK) substrate is colored in yellow, while its backbone is pink.

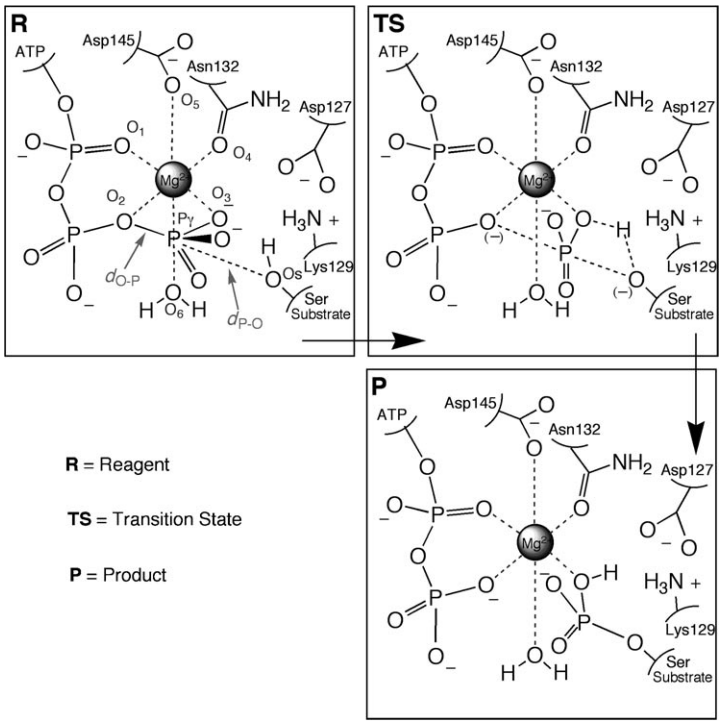


Figure 2. The S_N2-like single step mechanism for the γ-phosphate transfer catalyzed by CDKs, as postulated in reference [21], involves a nucleophilic attack of the serine oxygen (Ser-Os) on the γ-phosphorus (γP), along with a proton transfer from the Ser-Os atom to one oxygen of the γ-phosphate, coordinated to the Mg^{2+} ion.

potent and selective CDK inhibitors.[15,16] In this respect, a key step during the catalysis, carried out by protein kinases, is the deprotonation of the hydroxy group of the side chain of the substrate residue. This increases the nucleophilicity of the substrate and promotes the ATP γ-phosphate transfer. Because all protein kinases present a conserved Asp residue in the C-term domain (for instance, Asp127 in Cdk2, and Asp166 in cAMP-dependent protein kinase, PKA), this residue could play a key role in this step because of its vicinity to the substrate hydroxy group.[14] This structural evidence suggests a mechanism in which the Asp residue acts as a general base: Asp may receive one proton from the substrate hydroxy group, helping the formation of the final nucleophilic species, that is, the deprotonated Ser/Thr side chain. In agreement with this hypothesis, a recent computational investigation by Cheng et al.[17] supports a dissociative reaction mechanism for the phosphoryl transfer catalyzed by PKA. Here, Asp166 acts as a general base, thus enabling the deprotonation of the substrate serine. Other computational studies on the PKA kinase have reported a similar enzymatic mechanism.[18–20] In contrast, quantum-mechanical (QM) calculations performed on Cdk2 by some of us[21] indicate a phosphoryl transfer reaction with an associative character (Figure 2) and a direct proton transfer from the substrate hydroxy group of the serine side chain to the γ-phosphate of ATP; this proton transfer is concerted with the phosphoryl transfer and, hence, does not involve the Asp127 residue as a general base. Other computational and experimental studies support a similar associative mechanism, in agreement with this second proposed pathway.[22–25] In particular, computations carried out by Hart et al,[22,23] performed by means of semiempirical PM3 QM/MM calculations, suggest that the role of the conserved Asp residue as a general base is very unlikely. Also, considerations based on the relative basicities of the aspartate and serine suggest that the Asp residue is not likely to act as a general base because of its lower pK_a value.[26]

Here, we report a multistep computational study of the Cdk2 catalytic activity aimed at clarifying key aspects of the CDK reaction mechanism. The study is constituted of three steps. 1) DFT calculations on models of the catalytic residues. These calculations point to two plausible protonation states of the active site. 2) Molecular dynamics (MD) simulations of the two identified best protomers, and their equilibration in aqueous solution. 3) Constrained Car–Parrinello[27] (CP) QM/MM calculations,[28–30] performed to determine possible reaction pathways of the equilibrated protomers. The CP-QM/MM calculations allow the inclusion of the protein environment contributions (i.e., electrostatics) to the study of the reaction pathways. The latter point represents an important improvement to our previous QM study.[21]

Although many aspects of the reaction mechanism herein investigated resemble those found in our previous QM study, the calculated free energy barrier for the transfer of the γ-phosphate calculated by QM/MM simulations is considerably lower ($\approx$42 kcal mol^{-1} QM vs. $\approx$24 kcal mol^{-1} QM/MM). This supports the idea that a crucial role is played by the electrostatics of the protein environment and of the solvent waters in speeding up the catalysis. Overall, we cannot rule out the possibility that Asp127 acts as a general base in Cdk2, as seemingly indicated in PKA.[17] However, our calculations tend to indicate that Asp127 is important only for the stabilization of the near-attack conformation of the substrate serine, while it does not act as a general base. Clearly, our findings are not conclusive, and further

computational and experimental investigations are still required in order to promote a general consensus on the reaction mechanism catalyzed by protein kinases.

Results and Discussion

Here, we present and discuss our results concerning the ATP γ-phosphate transfer reaction mechanism, catalyzed by the Cdk2 enzyme. In detail, we discuss: 1) determination of the favored protonation states (protomers) of key residues in the catalytic site region; 2) equilibration of the identified best protomer and subsequent constrained CP-QM/MM simulations carried out in order to define possible pathways for the enzymatic reaction herein investigated.

Protonation states analysis: These calculations focused on the protonation state of: 1) The Asp127 residue, which was considered either in its protonated (protomer 1, P1) or deprotonated (protomer 2, P2) form (Figure 3). The protonation state of Asp127 is particularly interesting in consideration of its possible role as a general base in the deprotonation of the substrate hydroxy group. In fact, the Asp residue could receive a proton from the hydroxy group of the attacking nucleophile, leading to a dissociative mechanism for the subsequent enzymatic γ-phosphate transfer. 2) The ATP triphosphate moiety. In particular, we considered the protomers with an additional hydrogen atom either on one oxygen atom of the β-phosphate (protomer 3, P3) or on one oxygen atom of the γ-phosphate (protomer 4, P4) (Figure 3). The presence of this additional proton could lead to a possible dissociative mechanism for the phosphoryl transfer, in which the auto-dissociation of the γ-PO_3^- group would happen as follows: an intramolecular proton transfer, between the γ- and β-phosphate oxygen atoms, would cause a concomitant breaking of the O–γP bond.

Both the P1- and P2-optimized structures were in good agreement with the crystal structure, showing reasonable geometrical determinants (RMSD ≈0.7 Å in both protomers). The octahedral coordination shell of the Mg^{2+} cation was maintained as in the X-ray structure (Figure 3). Thus, both protomers were considered for subsequent MD simulations.

Conversely, the Mg^{2+} hexa-coordination was lost in both P3 and P4, with the consequent loss of the Mg^{2+}-centered octahedral geometry: the formation of one ATP intramolecular hydrogen bond, formed by the added hydrogen atom, caused a strong distortion around the Mg^{2+} ion and thus a

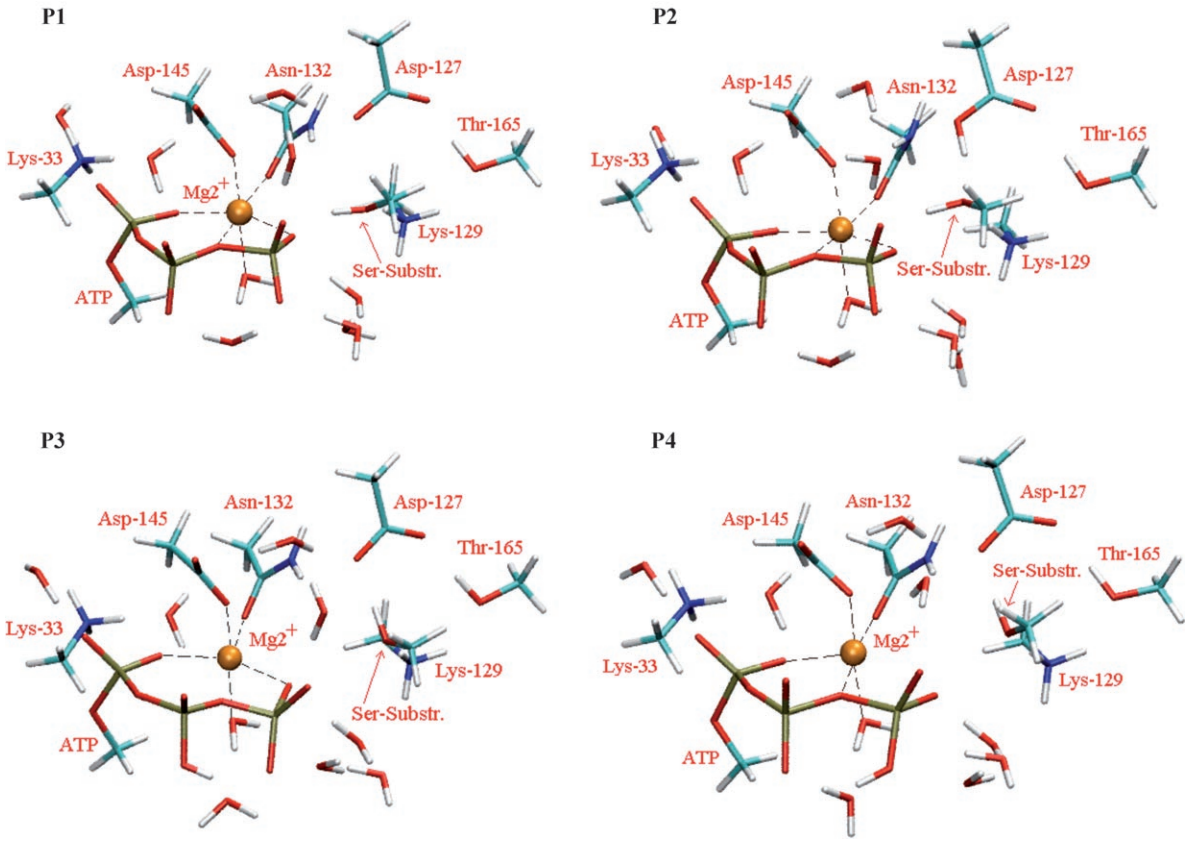
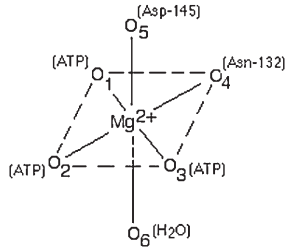


Figure 3. Optimized structures of protomers P1–P4. P1 presents Asp 127 in a deprotonated form, while P2 is in the protonated form. P3 and P4 have one protonated oxygen atom on the β- and γ- phosphate of ATP, respectively. While in P1 and P2 the hexa-coordination shell centered on Mg^{2+} is well maintained, both P3 and P4 optimizations led to a distorted model with consequent loss of the octahedral coordination centered on the Mg^{2+} ion (see also Table 1). Moreover, P3 and P4 showed an unfavorable position of the serine, far away from the γ-phosphate.

change in its coordination shell (Figure 3 and Table 1). As a result, the RMSDs were much larger than those of P1 and P2 ($\approx$1.8 Å and $\approx$1.6 Å for P3 and P4, respectively). This result is consistent with previous experimental findings indicating that the dominant form of the triphosphate tail of the $(Mg^{2+})MTP^{2-}$ complex is deprotonated under neutral conditions in solution.[31,32] Hence, these protomers (P3 and P4) were not taken into account for mechanistic studies.

Table 1. Mg–O bond lengths (Å) in the P1, P2, P3, P4 optimized structures. Comparison is made with the correspondent bond lengths of the X-ray structure.[14]



	X-ray	P1	P2	P3	P4
$Mg–O_1$	2.07	2.04	2.02	2.06	2.03
$Mg–O_2$	2.12	2.33	2.36	2.74	2.28
$Mg–O_3$	2.25	2.09	2.11	2.09	2.81
$Mg–O_4$	2.00	2.16	2.17	2.12	2.08
$Mg–O_5$	1.91	2.09	2.12	2.06	2.06
$Mg–O_6$	2.14	2.30	2.25	2.31	2.25

Reaction pathway: After geometry optimization of all protomers (P1, P2, P3, and P4), we performed a $\approx$3 ns MD run on the model system (MS) with P1 (MSP1) (i.e., Asp127 deprotonated), which was constituted by the entire Cdk2/cyclin A-substrate complex, solvated in a water box. The system was already equilibrated after about 1.2 ns, as shown by a plot of the RMSD of the protein as a function of simulated time (Figure 4). The Cdk2 fold turned out to be conserved, as well as the position of the substrate peptide, which remained near to the γ-phosphate ($\approx$3.45 Å average distance ($d_{P–O}$) between the ATP γ-phosphate atom, γP, and the nucleophilic oxygen atom of the serine side chain, Ser-Os). Asp127 interacted with Lys129 throughout the entire simulation, forming a stable salt-bridge. A configuration from the equilibrated simulations was taken after about 2.5 ns of dynamics, in which $d_{P–O}\approx$3.3 Å, and was used to initiate the QM/MM calculations.

After about 2 ps of unconstrained QM/MM dynamics, an estimate of the free energy profile as a function of $d_{P–O}$ (i.e., the reaction coordinate, RC) was computed by using constrained dynamics, as explained in the computational methods section. Constrained CP-QM/MM simulations of about 2–2.5 ps were performed for ten points of different RC values, from 3.60 Å to 2.15 Å. The shape of the resulting free energy profile was characterized by two minima: the reactant state (**Reag**) constituting the Michaelis–Menten complex and the final product (**Prod**) separated by a single TS

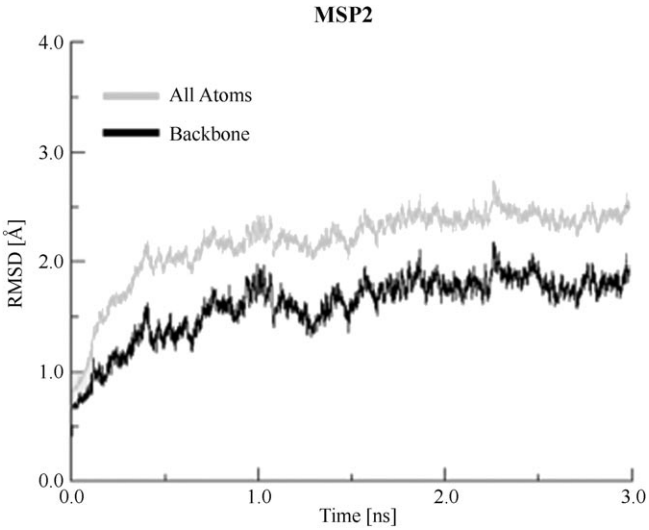


Figure 4. RMSD of all atoms or backbone atoms for MSP1 plotted as a function of simulated time.

maximum (**TS**). Simulations of the initial reactant state **Reag** (RC = 3.60 Å) showed a well-structured hydrogen bond network that stabilized the Mg^{2+}-centered catalytic site. The bond in breaking (i.e., $d_{O–P}$ indicated in Figure 2) was stable, showing an average value of about 1.68 Å. The salt-bridge, constituted by Asp127 and Lys129, was also stable along every step of the dynamics. This contributed to the maintaining of a good geometry of the enzymatic complex, especially in keeping the reciprocal positions of the reactants, formed by the terminal part of the ATP triphosphate moiety and the side chain of the nucleophilic serine residue. Furthermore, the Mg^{2+} hexa-coordination shell was well conserved along the entire path (Table 2).

Table 2. Selected average bond lengths (in Å) along the MSP1 reaction profile: 1) The six distances constituting the hexa-coordination shell centered on the Mg^{2+} ion and 2) the average length of forming and breaking bonds are reported.

	MSP1		
	Reagents	TS	Products
$Mg–O_1$	2.08 (0.05)	2.10 (0.05)	2.05 (0.05)
$Mg–O_2$	2.08 (0.05)	2.15 (0.1)	2.03 (0.05)
$Mg–O_3$	2.10 (0.05)	2.13 (0.05)	2.14 (0.05)
$Mg–O_4$	2.12 (0.05)	2.09 (0.05)	2.16 (0.05)
$Mg–O_5$	2.00 (0.05)	2.01 (0.05)	2.01 (0.05)
$Mg–O_6$	2.16 (0.05)	2.07 (0.05)	2.15 (0.05)
$d_{(P–O)}$	3.60	2.27	1.71 (0.05)
$d_{(O–P)}$	1.68 (0.05)	1.98 (0.1)	2.45 (0.1)

At RC = 2.27 Å, the transition state (**TS**) was reached: the average force acting on the constraint was essentially equal to zero. We observed the formation of the planar γ-phosphate geometry, which resembles the typical S_N2-like geometry for the phosphoryl transfer reaction mechanism: the nucleophile Ser-Os approached the γ-P atom on the opposite face of the γ-phosphate bond in breaking. The short length

of the bond in formation and in breaking (2.27 Å and 1.98 Å, respectively) suggested that this was an associative phosphoryl transfer mechanism, which resembled well the crystal structures of TS analogs,[24,25] and was in excellent agreement with the QM study of the same enzymatic reaction carried out using a different QM approach.[21]

During the catalytic pathway, Asp127 was constantly involved in a salt-bridge with Lys129, and it remained too far from the nucleophile Ser-Os atom to catalyze its deprotonation: the distance between the carboxylate oxygen atoms of Asp127 and the Ser-Os atom was stable at about 3.5 Å. In addition, a hydrogen bond between Ser-Os and the Asp127 carboxyl group was never formed along the reaction path. Test calculations, in which Asp127 was also included in the QM region, indeed showed that there was never a spontaneous deprotonation of Ser-Os during the investigated pathways, within the limited sample herein performed. To further address this issue, we considered models in which the proton was manually removed from the Ser-Os atom and added to Asp127. However, in this case the proton returned either to one oxygen of the ATP γ-phosphate or to the Ser-Os atom, within a few picoseconds of free QM/MM dynamics.

Thus, the structure presenting the protonated ATP γ-phosphate constitutes the most reliable TS geometry (**TS**, Figure 5), in very good agreement with the QM result. Also, the TS geometry agrees well with previous semi-empirical QM/MM calculations[22,23] performed by Hart et al., which proposed an alternative role for the deprotonated Asp in stabilizing the protonated form of the γ-phosphate in the catalytic site. This would induce TS-stabilization, while ruling out the role of the conserved Asp as a general base. However, because of the limited sampling of our study, and the limitations associated with the use of only one RC,[33] we cannot rule out the involvement of Asp127 as a general base in the catalysis. The substrate deprotonation remains an issue to be clarified. In this regard, it is important to notice that our calculations did not include solvation waters in the QM system. This might represent a major drawback in our simulations. In fact, as recently shown,[34–37] the inclusion of solvation waters in the QM system can facilitate proton transfers that use waters to perform efficient proton shuttles. This favors the formation of a good attacking and/or leaving group. Calculations including a larger QM region are currently in progress.

At RC = 2.15 Å, the average constraint force changed sign, indicating that the system is evolving towards the products. Thus, upon removal of the constraint, the system reached the product conformation, in which the ATP γ-phosphate was completely transferred to the Ser-Os, while ADP was formed (**Prod**, Figure 5). In detail, the products presented an average d_{P-O} of about 1.71 Å, and d_{O-P} was equal to about 2.45 Å (Table 2). Besides, the Mg^{2+} coordination was still well maintained, further supporting the catalytic importance of the chelating effect of the positive charged metal ion in the catalysis.

The reaction free energy barrier determined for the pathway **Reag** → **TS** → **Prod** was 23.7 ± 4.5 kcal mol^{-1} (see also the Supporting Information). This energy value is much lower than the activation energy of about 42 kcal mol^{-1} found for the same enzymatic reaction with a full QM approach using a smaller protein model system.[21] Indeed, our results showed a reaction mechanism consistent with that found in our previous QM study (a rather associative S_N2-like reaction mechanism). On the other hand, the decrease of the energy barrier supports the important role of the electrostatics of the protein frame and solvation waters during the enzymatic reaction.

As for previous computational studies of phosphoryl transfer in HIV-1 integrase,[38] Cdc42-GAP,[39] and sEH phosphatase activity,[34,35] no pentavalent phosphorane intermediate such as that observed in β-phosphoglucomutase[40–44] was observed. Although such an intermediate cannot be ruled out due to the flat nature of the TS region and the approximations of the calculations performed here,[33] it would significantly affect neither the reaction profile nor the reaction rate[45] since its energy could only be slightly lower (≈1–2 kcal mol^{-1}) than that of the surrounding TSs.

Whilst our previous QM study uses a localized Gaussian basis set (LBS), the present CP approach, used for the QM part of our QM/MM calculations, uses planewaves (PW). To verify the consistency between PW and LBS DFT calculations, we performed DFT/BLYP single point calculations on the geometries of the reagent and TS of the QM study using the CP method. Excellent agreement was found between the energies achieved with the two different methods (42.20 kcal mol^{-1} with LBS calculations vs. 42.19 kcal mol^{-1} with PW calculations). This result validates a direct comparison of the energies of the two studies, and provides further evidence of the barrier lowering observed in the present QM/MM study.

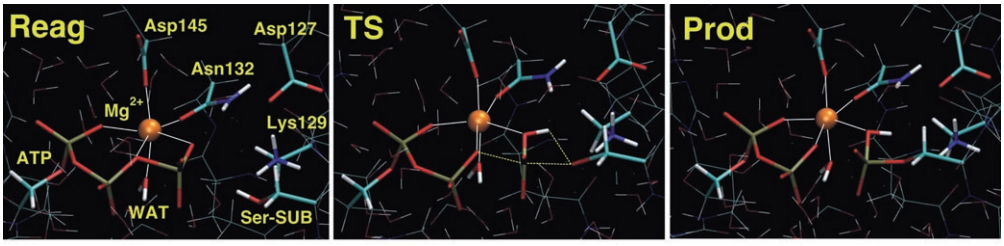


Figure 5. Representative snapshots of reagent (**Reag**), transition state (**TS**), and product (**Prod**) geometries along the pathway of MSP1.

A comparison of the enzymatic barrier with that of the phosphoryl transfer reaction in water is also useful in quantifying the catalytic power of the Cdk2 enzymatic mechanism herein proposed. For our case, the study of Akola and Jones[46] is the most relevant one since it concerns the hydrolysis of the Mg-complexed methyl triphosphate in water, studied by means of CPMD calculations. The barriers found for different mechanisms and reaction coordinates are about 35–40 kcal mol^{-1}, which further show the efficiency of the enzymatic mechanism herein proposed.

The second protomer (MSP2), in which Asp127 is protonated, was also investigated by means of constrained CP-QM/MM calculations after about 3 ns of equilibration performed by a classical MD run. We followed the same protocol of the MSP1 system (see the Supporting Information). As already mentioned, MSP2 differed from MSP1 only for the Asp127 form. Constrained CP-QM/MM simulations of about 2.0–2.5 ps were performed for seven points of different RC values, from 3.39 Å to 2.43 Å. All the details about the procedure of the calculations were the same as discussed previously for MSP1.

In this case, the near-attack conformation was strongly distorted compared to that of MSP1. The protonation of Asp127 led to a different hydrogen bond network in the catalytic site, which was mainly evident from the loss of the salt-bridge previously formed by Asp127 and Lys129. As a result, Ser-Os attacked the γ-phosphate group on the same side of the bond in breaking (see also the Supporting Information). Clearly, this was a highly unfavorable approach, which led to an unphysically large activation free energy (36 kcal mol^{-1}) at the last calculated point (LCP) along the reaction path (see Figure 4 in the Supporting Information). After LCP, the system was no longer stable, forcing the termination of the simulations along this pathway. In summary, our calculations ruled out the P2 protomer, further substantiating the fact that Asp127 has to be deprotonated in the Cdk2 active form.

Conclusion

We have presented a computational investigation of the phosphoryl transfer reaction catalyzed by the Cdk2 enzyme. The favored enzymatic pathway is an associative S_N2-like phosphoryl transfer mechanism, in agreement with a previous QM study of the same enzymatic reaction.[21] This resembles well the crystal structures of transition state analogs[24,25] and it is associated with a free energy barrier (≈24 kcal mol^{-1}) much lower than the activation barrier obtained in the QM study (≈42 kcal mol^{-1}), indicating the key role of the protein frame.

The deprotonation of the hydroxy group of the side chain of the substrate residue is a requisite chemical step during the catalysis carried out by protein kinases. This increases the nucleophilicity of the substrate and promotes the ATP γ-phosphate transfer. Importantly, all protein kinases present a conserved Asp residue in the C-term domain (Asp127 in Cdk2), which is indicated as a possible key residue in the catalysis due to its vicinity to the substrate hydroxy group.

Our calculations clearly indicate Asp127 to be deprotonated in the active form. However, our findings do not support its role as a general base during the reaction mechanism, while Asp127 is essential in conserving the near-attack conformation of the substrate residue. In fact, we did not observe any spontaneous deprotonation of the substrate serine in our simulations. This might be at the origin of the relatively high activation free energy barrier herein reported. However, our findings show that the more stable TS geometry is the one having the proton of the substrate hydroxy group transferred on the ATP γ-phosphate, while the Asp127 remains far from the reactive region. Importantly, these results agree well with previous semiempirical calculations,[22,23] which do not support the role of the conserved Asp as a general base. Nevertheless, further investigations on possible pathways for the substrate deprotonation are needed, since the solvation waters might be important for the necessary deprotonation of the substrate and a subsequent water-mediated proton shuttle. Finally, because of the high amino acid sequence conservation among the whole family of CDKs, we suggest a generalization of the results found here for the CDK protein family. Also, this information might be useful in designing new and more specific TS analogues as CDK inhibitors.

Computational Methods

Structural models: Every model was constructed from the 2.2 Å-resolution X-ray crystallographic structure of the Michaelis complex formed by the pCdk2/cyclin A complex and a substrate analogue peptide (PDB entry: 1QMZ).[14] The structure contains the entire pCdk2 protein (298 residues), the inactive AMPPNP then replaced by ATP, an Mg^{2+} ion, residues 174–432 of cyclin A (which contain the region binding to Cdk2), and the HHA**S**PRK substrate (Sub) where **S** is the serine which is phosphorylated by ATP during the enzymatic catalysis. The presence of AMPPNP then replaced by ATP, a Mg^{2+} ion, an optimal substrate analog, and the phosphorylated Thr160 made a reliable active conformation from this enzymatic pCdk2/cyclin A/Sub complex.

Models for QM calculations: Four large and different model systems (MSs) were considered as possible protonation states. Every MS included: ATP, Asp127, Asp145, Asn132, Lys33, Lys129, Thr165, and the serine substrate, replaced respectively by a methyl triphosphate moiety, two acetate molecules, an acetamide, two methyl amines, and two methyl alcohol molecules. Nine crystallographic water molecules were also included in the final MSs along with the catalytic Mg^{2+} ion. MSs of protomer 1 (P1) and protomer 2 (P2) resembled the catalytic site including either the deprotonated or protonated Asp127 form, respectively. Protomer 3 (P3) and protomer 4 (P4) included the protonated ATP triphosphate moiety: P3 was protonated on one oxygen atom of the β-phosphate of ATP, while P4 was protonated on one oxygen atom of the γ-phosphate of ATP. Each MS constituted a total of 96 (deprotonated MS, P1) or 97 (protonated MSs, P2, P3, P4) atoms. The optimized structures of each MS are shown in Figure 3.

Models for classical MD and QM/MM calculations: Two model systems of the pCdk2/cyclin A/Sub complex presenting the two favorite protonation states (i.e., protomers P1 and P2) were considered. Model System 1 presented the protonation state of P1 (MSP1), while Model System 2 presented the protonation state of P2 (MSP2). Hence, MSP1 and MSP2 differed only in the protonation state of Asp127. The 382 crystallographic

water molecules were retained. Then, both MSP1 and MSP2 were immersed in a box ($82.1\times89.2\times100.9$ Å) containing about 17000 water molecules. The total number of atoms was about 62000.

Both MSP1 and MSP2 were equilibrated by means of about 3 ns of classic MD calculations. The AMBER force field[47] was adopted for the simulations, according to ref. [12]. Details of the setup of classical MD simulations are reported in the Supporting Information.

QM calculations—protonation state analysis: Density functional theory (DFT) calculations were carried out using the Becke and Lee-Yang Parr (BLYP) exchange and correlation functional.[48,49] The Kohn–Sham orbitals were expanded in basis set plane-waves up to a cutoff of 70 Ry. The interactions between electrons and ionic cores were described by Martins–Troullier norm-conserving pseudo-potentials.[50] A $20\times16\times17$ Å cell was used for every protonation state model system (i.e., P1, P2, P3, and P4), treated as isolated. Geometry optimizations were performed using the direct inversion in the iterative subspace (DIIS) method until a convergence of 0.0005 au on the gradient was reached.

To discriminate between the four protomers considered, we looked at the following geometrical features of the final optimized structures: 1) maintenance of the Mg^{2+} hexa-coordination shell geometry of the starting crystal structure; 2) RMS deviations calculated by superposition of protein heavy atoms of optimized models and crystal structures.

Car–Parrinello QM/MM calculations: Both MSP1 and MSP2 were partitioned into two regions: one treated at the QM level and the other treated classically (MM). The QM part was treated using DFT-based molecular dynamics simulations, following the approach of Car and Parrinello,[27] (CP). The MM part was treated with the AMBER force field.[47] In both cases, the QM region included 32 atoms, specifically: ATP methyl triphosphate moiety, Mg^{2+} ion, side chain of serine substrate, and the final part of the Lys129 side chain. DFT calculations of the QM region were carried out using the BLYP exchange and correlation functional.[48,49] Unfortunately, the hybrid B3LYP functional is still too demanding for these costly calculations and a tradeoff between accuracy and system size was needed. However, the DFT/BLYP CP approach has been shown to describe accurately a variety of enzymatic systems and DNA complexes.[33,35,51–53] Moreover, the use of first principle CP dynamics includes entropic effects, which is an important aspect in the determination of the free energy barriers of enzymatic reactions.

The C_4–C_5 (ATP), the C_δ–C_ε (Lys129), and the C_α–C_β (Ser-Substrate) bonds crossed the QM/MM interface. In order to saturate the chemical bond of ATP-C_5, Lys129-C_ε, and Ser-Substrate-C_β, parameterized monovalent pseudopotentials of ATP-C_4, Lys129-C_δ, and Ser-Substrate-C_α were used.[54] Both QM regions of MSP1 and MSP2 used a supercell of $18\times15.3\times15.3$ Å, treated as isolated systems. Further features of the QM part treatment are the same as described in the QM Computational Methods section.

QM/MM simulations were carried out at an average temperature of 300 K, using a Nosé–Hoover thermostat[55,56] of 500 cm^{-1} frequency. A time step of 0.121 fs and a fictitious electron mass of 900 au were used. Electrostatic interactions between quantum and classical regions of the systems were described within a fully Hamiltonian coupling scheme.[54] Also, a rigorous treatment of the electrostatic interaction between QM and MM regions was implemented as reported in reference [57].

The overall procedure of our QM/MM calculations included an initial equilibration step of the MD, as starting snapshots of both MSP1 and MSP2. In detail, we performed a first step where only the MM part was free to move for a few hundred steps, while the QM part was frozen. Then, the whole system was allowed to move and heated up to 300 K (≈2 ps of free QM/MM dynamics). Finally, a good snapshot during the equilibrated simulations was chosen for performing constrained QM/MM calculations.

To follow the phosphoryl transfer reaction profile, a suitable reaction coordinate (RC) was chosen. To this end, the distance between the ATP γ-phosphorus atom (γP) and the nucleophilic oxygen atom of the serine substrate side chain (Os) was used as the main RC (i.e., distance γP–Os = $d_{P\text{-}O}$) for the enzymatic reaction under study (Figure 2). So-called blue-moon ensemble simulations were carried out for these systems constrained at different values of the RC, leaving all other degrees of freedom free to evolve. The protocol for constrained QM/MM simulations along the chosen RC for both systems was as follows: for MSP1, from 3.60 Å (reagent) to 2.15 Å, in ten steps (d_{P-O} = 3.60, 3.28, 3.12, 2.91, 2.77, 2.64, 2.54, 2.43, 2.27, and 2.15 Å); for MSP2, from 3.39 Å (reagent) to 2.43 Å in seven steps (d_{P-O} = 3.39, 3.12, 2.96, 2.75, 2.61, 2.53, and 2.43 Å). Each point was simulated for about 2–2.5 ps, of which the first at about 0.5–0.8 ps were used for equilibration. Considering both models, a total of about 50 ps of constrained QM/MM dynamics were performed. The activation free energy was calculated by using thermodynamic integration.[58] The estimated energy, which employs costly computational calculations, should be considered approximate. For a more accurate estimation of the enzymatic activation free energies, several independent paths should be considered, which is currently beyond the available computational resources.

Acknowledgements

This work was funded by the INFM project and all calculations were performed at the supercomputer center CINECA (Bologna, Italy). PRINCOFIN is also gratefully acknowledged for financial support.

[1] D. O. Morgan, *Annu. Rev. Cell Dev. Biol.* **1997**, *13*, 261–291.
[2] P. D. Jeffrey, A. A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massague, N. P. Pavletich, *Nature* **1995**, *376*, 313–320.
[3] N. P. Pavletich, *J. Mol. Biol.* **1999**, *287*, 821–828.
[4] F. H. Espinoza, A. Farrell, H. Erdjument-Bromage, P. Tempst, D. O. Morgan, *Science* **1996**, *273*, 1714–1717.
[5] D. O. Morgan, *Nature* **1995**, *374*, 131–134.
[6] J. W. Harper, P. D. Adams, *Chem. Rev.* **2001**, *101*, 2511–2526.
[7] L. N. Johnson, R. J. Lewis, *Chem. Rev.* **2001**, *101*, 2209–2242.
[8] J. F. Morrison, E. Heyde, *Annu. Rev. Biochem.* **1972**, *41*, 29–54.
[9] H. L. De Bondt, J. Rosenblatt, J. Jancarik, H. D. Jones, D. O. Morgan, S. H. Kim, *Nature* **1993**, *363*, 595–602.
[10] P. Cohen, *Nat. Rev. Drug Discovery* **2002**, *1*, 309–315.
[11] J. Dancey, E. A. Sausville, *Nat. Rev. Drug Discovery* **2003**, *2*, 296–313.
[12] M. De Vivo, A. Cavalli, G. Bottegoni, P. Carloni, M. Recanatini, *Proteins* **2006**, *62*, 89–98.
[13] N. R. Brown, M. E. Noble, A. M. Lawrie, M. C. Morris, P. Tunnah, G. Divita, L. N. Johnson, J. A. Endicott, *J. Biol. Chem.* **1999**, *274*, 8746–8756.
[14] N. R. Brown, M. E. Noble, J. A. Endicott, L. N. Johnson, *Nat. Cell Biol.* **1999**, *1*, 438–443.
[15] M. E. Noble, J. A. Endicott, L. N. Johnson, *Science* **2004**, *303*, 1800–1805.
[16] A. Cavalli, P. Carloni, M. Recanatini, *Chem. Rev.* **2006**, *106*, 3497–3519.
[17] Y. Cheng, Y. Zhang, J. A. McCammon, *J. Am. Chem. Soc.* **2005**, *127*, 1553–1562.
[18] N. Diaz, M. J. Field, *J. Am. Chem. Soc.* **2004**, *126*, 529–542.
[19] G. Henkelman, M. X. LaBute, C. S. Tung, P. W. Fenimore, B. H. McMahon, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15347–15351.
[20] M. Valiev, R. Kawai, J. A. Adams, J. H. Weare, *J. Am. Chem. Soc.* **2003**, *125*, 9926–9927.
[21] A. Cavalli, M. De Vivo, M. Recanatini, *Chem. Commun.* **2003**, 1308–1309.
[22] J. C. Hart, D. W. Sheppard, I. H. Hillier, N. A. Burton, *Chem. Commun.* **1999**, 79–80.
[23] J. C. Hart, I. H. Hillier, N. A. Burton, D. W. Sheppard, *J. Am. Chem. Soc.* **1998**, *120*, 13535–13536.
[24] I. Schlichting, J. Reinstein, *Biochemistry* **1997**, *36*, 9290–9296.
[25] I. Schlichting, J. Reinstein, *Nat. Struct. Biol.* **1999**, *6*, 721–723.
[26] J. Zhou, J. A. Adams, *Biochemistry* **1997**, *36*, 2977–2984.
[27] R. Car, M. Parrinello, *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

[28] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227–249.
[29] M. J. Field, P. A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700–733.
[30] J. Gao, D. G. Truhlar, *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.
[31] H. Sigel, R. Tribolet, R. Malinibalakrishnan, R. B. Martin, *Inorg. Chem.* **1987**, *26*, 2149–2157.
[32] R. Tribolet, H. Sigel, *Eur. J. Biochem.* **1988**, *170*, 617–626.
[33] K. Spiegel, U. Rothlisberger, P. Carloni, *J. Phys. Chem. B* **2006**, *110*, 3647–3660.
[34] M. De Vivo, B. Ensing, M. Dal Peraro, G. A. Gomez, D. W. Christianson, M. L. Klein, *J. Am. Chem. Soc.* **2007**, *129*, 387–394.
[35] M. De Vivo, B. Ensing, M. L. Klein, *J. Am. Chem. Soc.* **2005**, *127*, 11226–11227.
[36] F. L. Gervasio, V. Schettino, S. Mangani, M. Krack, P. Carloni, M. Parrinello, *J. Phys. Chem. B* **2003**, *107*, 6886–6892.
[37] A. Laio, F. L. Gervasio, J. VandeVondele, M. Sulpizi, U. Rothlisberger, *J. Phys. Chem. B* **2004**, *108*, 7963–7968.
[38] F. Bernardi, A. Bottoni, M. De Vivo, M. Garavelli, G. Keseru, G. Naray-Szabo, *Chem. Phys. Lett.* **2002**, *362*, 1–7.
[39] A. Cavalli, P. Carloni, *J. Am. Chem. Soc.* **2002**, *124*, 3763–3768.
[40] K. N. Allen, D. Dunaway-Mariano, *Science* **2003**, *301*, 5637.
[41] G. M. Blackburn, N. H. Williams, S. J. Gamblin, S. J. Smerdon, *Science* **2003**, *301*, 1184; author reply 1184.
[42] R. R. Holmes, *Acc. Chem. Res.* **2004**, *37*, 746–753.
[43] S. D. Lahiri, G. Zhang, D. Dunaway-Mariano, K. N. Allen, *Science* **2003**, *299*, 2067–2071.
[44] L. W. Tremblay, G. Zhang, J. Dai, D. Dunaway-Mariano, K. N. Allen, *J. Am. Chem. Soc.* **2005**, *127*, 5298–5299.
[45] J. Florian, A. Warshel, *J. Phys. Chem. B* **1998**, *102*, 719–734.
[46] J. Akola, R. O. Jones, *J. Phys. Chem. B* **2003**, *107*, 11774–11783.
[47] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
[48] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
[49] C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[50] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, *43*, 1993–2006.
[51] P. Carloni, U. Rothlisberger, M. Parrinello, *Acc. Chem. Res.* **2002**, *35*, 455–464.
[52] M. Dal Peraro, L. I. Llarrull, U. Rothlisberger, A. J. Vila, P. Carloni, *J. Am. Chem. Soc.* **2004**, *126*, 12661–12668.
[53] A. Magistrato, P. Ruggerone, K. Spiegel, P. Carloni, J. Reedijk, *J. Phys. Chem. B* **2006**, *110*, 3604–3613.
[54] A. Laio, J. VandeVondele, U. Rothlisberger, *J. Chem. Phys.* **2002**, *116*, 6941–6947.
[55] S. Nose, *J. Chem. Phys.* **1984**, *81*, 511–519.
[56] W. G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695–1697.
[57] A. Laio, J. VandeVondele, U. Rothlisberger, *J. Phys. Chem. B* **2002**, *106*, 7300–7307.
[58] G. Ciccotti, M. Ferrario, J. T. Hynes, R. Kapral, *Chem. Phys.* **1989**, *129*, 241–251.

Received: January 10, 2007
Revised: May 15, 2007
Published online: July 18, 2007

