J Comput Aided Mol Des (2007) 21:603–615
DOI 10.1007/s10822-007-9143-x

Comparative study of the prereactive protein kinase A Michaelis complex with Kemptide substrate

Manuel Montenegro · Mireia Garcia-Viloca · Àngels González-Lafont · José M. Lluch

Received: 7 June 2007 / Accepted: 23 October 2007 / Published online: 16 November 2007

Abstract In the present work we have modeled the Michaelis complex of the cyclic-Adenosine Monophosphate Dependent (cAMD) Protein Kinase A (PKA) with Mg_2ATP and the heptapeptide substrate Kemptide by classical molecular dynamics. The chosen synthetic substrate is relevant for its high efficiency and small size, and it has not been used in previous theoretical studies. The structural analysis of the data generated along the 6 ns simulation indicates that the modeled substrate–enzyme complex mimics the substrate binding pattern known for PKA. The values of the average prereactive distances obtained from the simulation do not exclude any of the two limiting situations proposed as mechanisms in the literature for the phosphorylation reaction (dissociative and associative) because the system oscillates between configurations compatible with each of them. Furthemore, the results obtained for the average interaction distances between active site residues concord in suggesting the plausability of an alternative third reaction mechanism.

Keywords Classical molecular dynamics · Kemptide · Molecular modeling · Protein kinase A · Reaction mechanism · Substrate binding

Electronic supplementary material The online version of this article (doi:10.1007/s10822-007-9143-x) contains supplementary material, which is available to authorized users.

M. Montenegro · M. Garcia-Viloca (✉) · À. González-Lafont · J. M. Lluch
Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
e-mail: mireia@bioinf.uab.es

À. González-Lafont · J. M. Lluch
Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Introduction

Protein kinases are highly widespread in nature, more or less 2% of eukaryotic genes encode them. These enzymes catalyze protein phosphorylation, a reversible covalent modification consistent in the transfer of the γ-phosphoryl group of ATP (adenosine triphosphate) to a specific protein's serine, threonine or tyrosine residue. This process is critical in many aspects of cell life (cellular regulation, metabolic pathways, gene transcription, signal transduction, etc) and it makes that the protein kinase family is involved in a large number of pathologies, including inflammatory disorders, autoimmune diseases and cancer. For this reason, there has been a major effort to understand the reaction pathway of protein kinases [1], and to design specific and potent inhibitors that might be useful as therapeutic agents [2, 3].

The protein kinase superfamily is large and structurally diverse, but all the enzymes of the group have a highly conserved catalytic core with a bilobal structure that suggests the same catalytic mechanism [4, 5]. In this work, we have selected protein kinase A (PKA), which is the best characterized member of the protein kinase family, as the object of our study. The inactive PKA holoenzyme is a tetramer composed of two regulatory and two catalytic subunits (C-subunits) that become dissociated after the two messenger cAMP (cyclic-adenosine monophosphate) molecules bind to the regulatory subunits. In Fig. 1 the crystallographic structure of the PKA catalytic subunit is depicted [6]. Once the two catalytic subunits become dissociated, they work independently.

PKA[1] is a multi-functional protein kinase that, in previous studies involving the phosphorylation of both protein

[1] From now on, with the name PKA we will refer solely to the catalytic subunit of the enzyme and not to the entire tetramer.

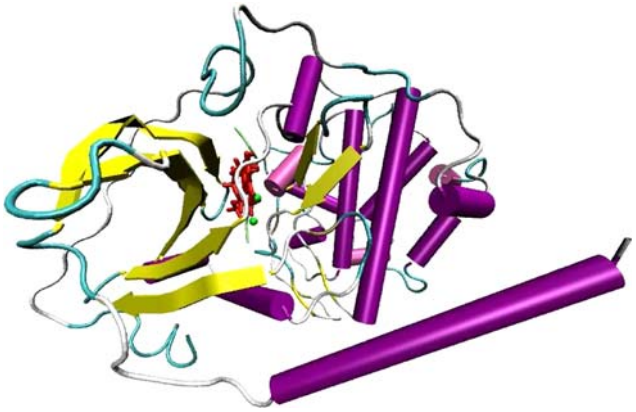

Fig. 1 Structure of the catalytic subunit of protein kinase A

and peptide substrates, has been shown to be an enzyme that recognizes a restricted region of the local sequence around the phosphorylation site [7–11]. This consensus sequence recognized by PKA is Arg-Arg-X-P-Z, where X is a small residue, P is the residue that undergoes the phosphorylation (a serine or a threonine) and Z a large hydrophobic one. The substrate used in the present theoretical study is the synthetic heptapeptide Leu-Arg-Arg-Ala-Ser-Leu-Gly (Kemptide) [12] that corresponds to the consensus sequence with an alanine as the small residue at the fourth position (X = Ala16), a serine (P = Ser17) at the phosphorylation site, a leucine as the large hydrophobic residue at the sixth position (Z = Leu18), and two extra residues, a leucine (Leu13) as the N-terminal residue and a glycine (Gly19) at the C-terminus of the heptapeptide. The thermodynamics and kinetics of the PKA–ATP–Kemptide complex have been experimentally analyzed by several groups [13, 14], and this makes it a suitable system to theoretically establish the most favourable catalytic pathway. The main conclusions of those experimental studies are that Kemptide is an excellent substrate for PKA with an apparent K_m that takes different values depending on the buffer conditions (31.1 ± 3 μM, 16 ± 0.9 μM, and 4.9 ± 1.4 μM) [9, 12, 14] and a remarkable phosphorylation rate (500 s^{-1}) [14]. K_m appears to depend on the presence of the two arginines (Arg14 and Arg15) at the N-terminal side of the phosphorylated serine [12]. These two arginines act as important substrate specificity determinants for PKA.

In Fig. 2 a schematic representation of the atomic interactions in the reactive core of the enzyme is shown. As described by previous experimental works [7, 12, 15] the substrate is anchored to the enzyme by interactions of Arg14 with Glu127 and Thr51, of Arg15 with Glu170 and Glu230, between Leu18 and Gly200, and finally, between Ser17 (the phosphorylation undergoing residue) and Lys168 (see Fig. 2a).

The role of the two Mg^{++} ions in the catalytic mechanism is not completely clear, but since the essential presence of the two atoms has been proved in other enzymes of the same family [16], we consider a system with two Mg^{++} ions in the active center (see Fig. 2b, c). Mg_1 seems to operate as a real catalyst facilitating the transfer of the γ-phosphoryl group from ATP to Ser17, whereas Mg_2 has been identified by some authors as an inhibitor. Both ions present octaedric coordination in the PKA active site.

There are also intriguing questions concerning the roles of Asp166 and Lys168 residues [8]. Asp166 is a highly conserved residue in all the members of the kinase superfamily and its key presence has been clearly proved by mutational studies. In several X-ray crystallographic structures [1, 6, 8] the carboxylate group of this residue was found near the hydroxyl proton of the peptide substrate, and this led to the early hypothesis that this residue may act as a general-base catalyst. However, the rate of phosphoryl transfer in PKA, measured using pre-steady-state kinetics, is neither subject to a solvent deuterium isotope effect nor pH dependent [17], which does not support the early proton transfer implied by conventional base catalysis. In view of the conflict, several experimental studies have been focused on the role of the conserved aspartate in the catalytic loop of all serine/threonine and tyrosine kinases. The current hypothesis are that Asp166 could either help to correctly orient the reactants or act as a proton acceptor late in the reaction process [1, 17, 18]. In addition, several research groups have carried out theoretical studies of the phosphorylation process catalyzed by PKA and have arrived to different conclusions concerning the reaction pathway and the role of the different enzyme residues [16, 19–26]. As a result of these numerous experimental and theoretical works two possibilities are drawn for the phosphoryl transfer step mechanism that are represented in Fig. 3. The first one (Scheme A) consists in the phosphoryl- and proton-transfers between the substrate and the ATP molecule, which is also known as the associative mechanism. The second proposed mechanism (Scheme B) is a dissociative path in which the substrate proton is accepted by the enzymatic residue Asp166 and the γ-phosphoryl group passes from ATP to the substrate. Early semiempirical quantum mechanical (QM) calculations on cluster models of the enzyme–ATP–substrate complex, or semiempirical quantum mechanical/molecular mechanical (QM/MM) calculations on complete models of the biological system gave results that support the mechanism in Scheme A [20, 22, 23]. More recently, the exploration of the potential energy surface for the phosphoryl transfer reaction has been carried out with higher QM levels of theory on cluster [16, 24] or complete [19, 25] models of the system. These later theoretical studies are consistent with the role of Asp166 as a proton-trap late in the reaction process.

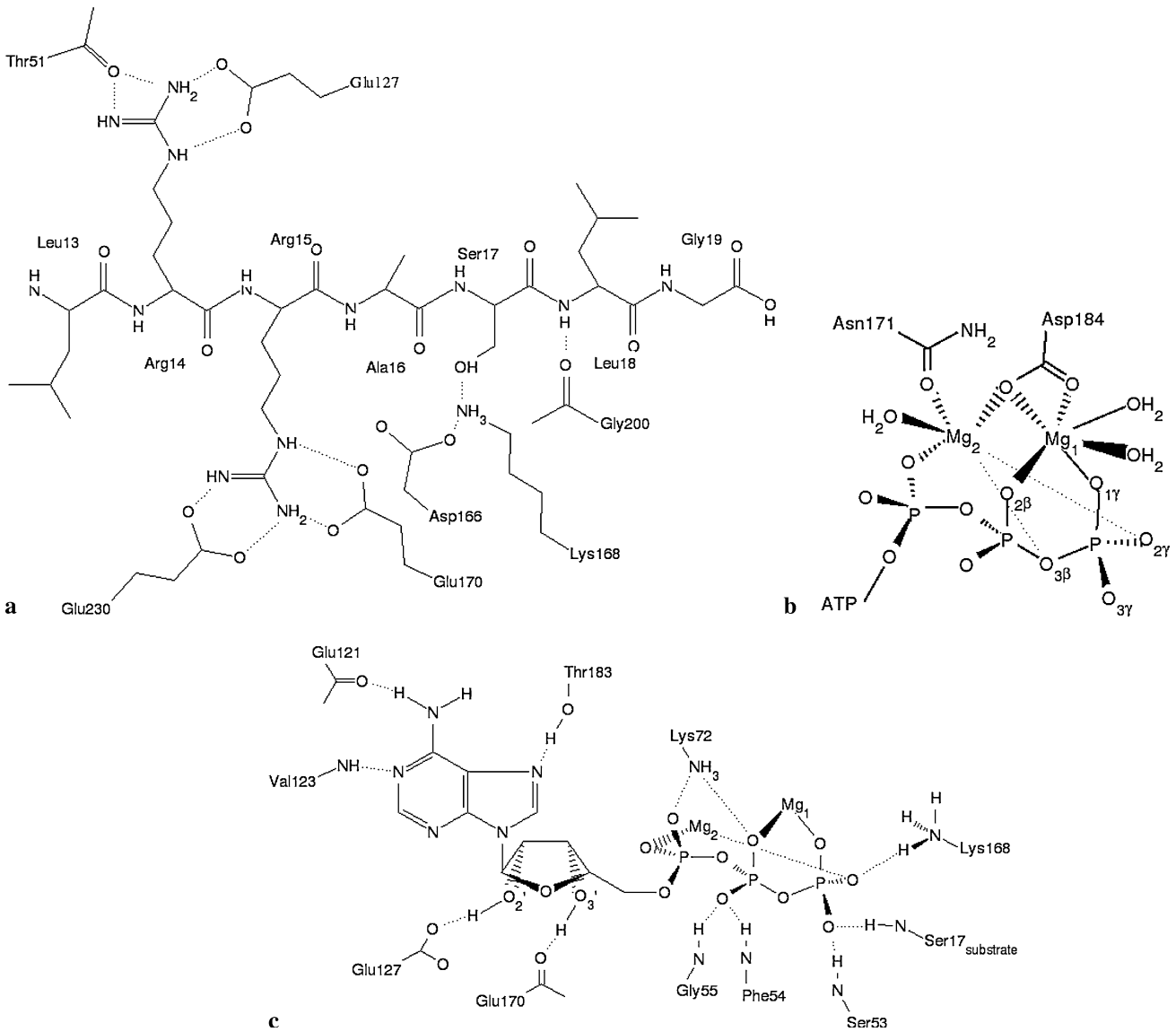

Fig. 2 Scheme of the intermolecular interactions: (**a**) PKA-Substrate; (**b**) Magnesium Ions; (**c**) ATP-PKA

On the other hand, the experimental results found in the literature indicate that the kinetic properties of the enzyme depend on the nature of the substrate [8, 12, 14, 17], but this issue has not been analyzed in the previous theoretical studies. Three of the most recent ones have analyzed the wild type PKA complex with 20-residue peptide substrate-analogs (model substrate and SP20 in Table 1), built up from the 20-residue peptide inhibitor PKI(5–24) [27] located in the kinase active site in the crystallographic structure [1, 6]. In the study by Diaz and Field [16] a MD simulation of $\sim$1.8 ns was carried out and the results support the proposal in which the interaction between the substrate Ser17 residue and the conserved Asp166 residue could promote phosphoryl transfer by limiting the number of hydroxyl rotamers and positioning the nucleophile for productive substitution on the γ-phosphoryl group of ATP. In addition, their DFT study of the catalytic reaction mechanism in a cluster model of the PKA active site, suggests that Asp166 can behave as the residue that accepts the proton delivered by the substrate peptide along a mostly dissociative phosphorylation reaction pathway. In two recent studies by Cheng et al. [19, 25], a QM/MM potential energy surface was explored for a complete enzyme–ATP–substrate complex and the results also indicated that the phosphorylation reaction catalyzed by PKA is mainly dissociative and that Asp166 serves as the catalytic base to accept the proton delivered by the substrate peptide. However, their results are not consistent with the second proposed role of Asp166 as a structural anchor to maintain the configuration. In addition, Cheng et al. [19] showed

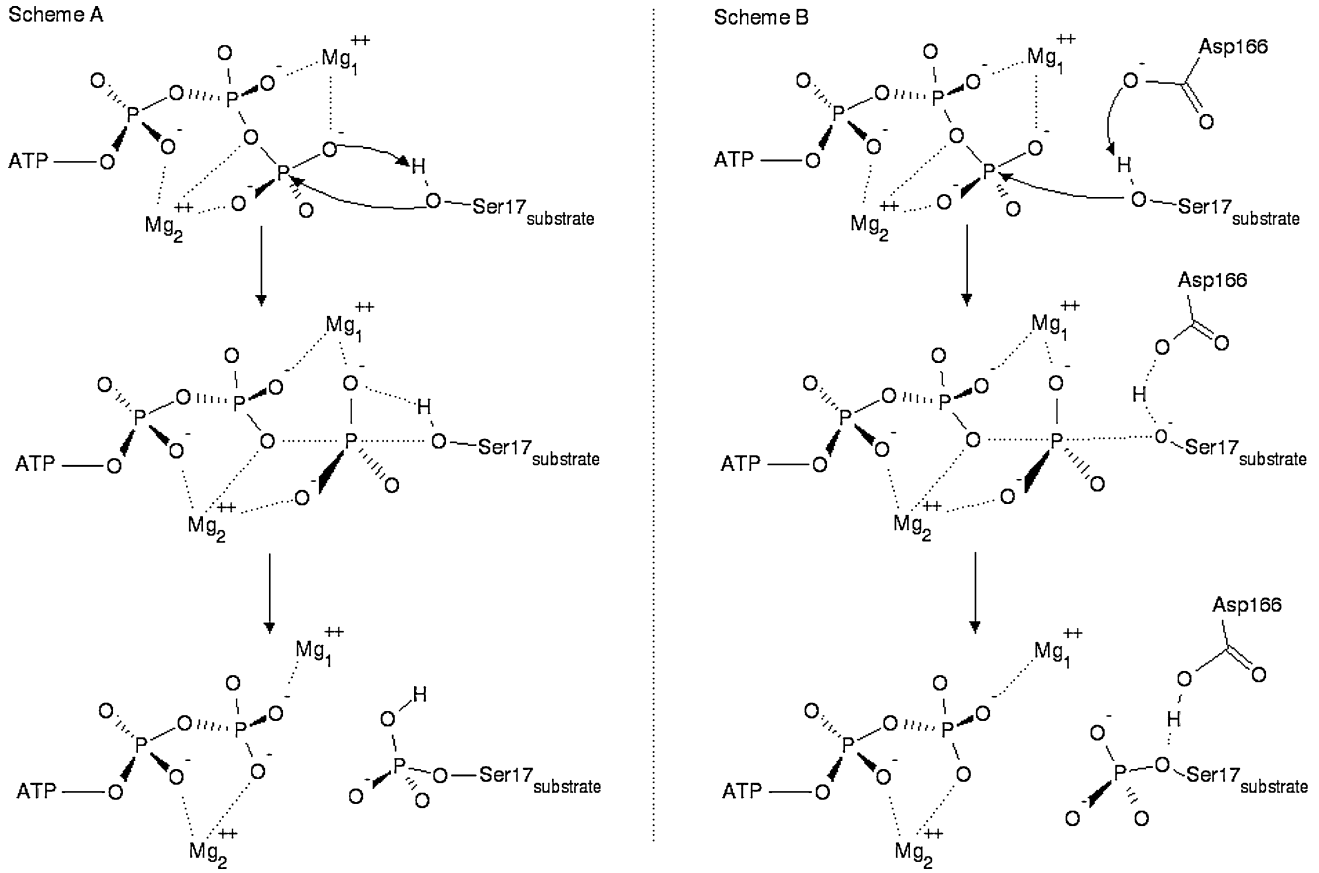

Fig. 3 Reaction scheme for the phosphoryl transfer step: Scheme A is the associative mechanism which consists in the phosphoryl- and proton-transfers between the substrate and the ATP molecule. Scheme B represents the dissociative reaction pathway; the proton of Ser17 passes to an oxygen of Asp166 during the nucleophilic displacement

that Lys168 serves to keep ATP and the substrate peptide in the near-attack reactive conformation.

The aim of the present work is to theoretically simulate for the first time the PKA–Mg_2ATP–Kemptide complex. Our choice of target substrate for PKA has been Kemptide because detailed thermodynamic and kinetic experimental analyses have been performed on the phosphorylation of this substrate and, indeed, the experimental results using this peptide and showing its high efficiency have been more consistent than with other substrates [10, 11, 14]. However, there is no structure of any kinase where a small substrate peptide is bound in the active site. Thus, the purpose of our MD simulations is to determine the dynamical stability of the Kemptide substrate and its most important prereactive contacts in the active site of PKA, which determine substrate specificity. In addition, this study contributes to further support that the minimum residue sequence (consensus sequence) Arg-Arg-Ala-Ser-Leu is recognized and bound in the PKA active site, which might provide insights into the mechanism-based design of potent kinase inhibitors.

Table 1 Inhibitor and substrates of PKA

	Amino acid sequence
Part of PKI	TDVE**TTYADFIASGRTGRRNAIHD**ILVS...
PKI(5–24)	TTYADFIASGRTGRRNAIHD
SP20	TTYADFIASGRTGRR**AS**IHD
Model substrate	TTYADFIASGRTGRRN**S**IHD
Kemptide	LRRASLG

PKI is the protein kinase inhibitor from rabbit skeletal muscle, the residues represented with bold characters are those corresponding to the inhibitor PKI(5–24). SP20 is the 20-residue peptide substrate used by various experimental and theoretical research groups, while the fourth reported peptide is a model substrate used by Diaz and Field. The residues represented with bold characters are the modified ones with respect to PKI(5–24). Kemptide is the synthetic hepta-peptide used in the present work

Computational details

Model of the enzyme–substrate–cofactor Michaelis complex

The starting coordinates of the enzyme's catalytic subunit were taken from the 2.0 Å resolution crystallographic

structure with PDB-ID-code 1CDK [6], which corresponds to the closed active conformation of the enzyme. The crystallographic structure contains the enzyme catalytic subunit, an ATP analogue (AMP-PNP), two Mn^{++} ions, a 20-residue peptide inhibitor (PKI(5–24)), and 149 water molecules. Our system was built changing the Mn^{++} ions to the biologically active Mg^{++} ions and the AMP-PNP molecule to ATP. In order to mantain the active configuration of PKA we kept at position 197 a phosphorylated threonine. The Kemptide was positioned in the active centre of PKA by manually modeling it from the PKI(5–24) inhibitor, which is bound to the enzyme in the X-ray structure. With respect to the inhibitor crystallographic structure, we have deleted all the residues from Thr1 to Thr12 and the last one, Asp20.[2] Then, we have modified the remaining residues to obtain the Kemptide substrate: we changed Gly13 to Leu13, Asn16 to Ala16, Ala17 to Ser17, Ile18 to Leu18, and His19 to Gly19.

The coordinates of enzyme and substrate hydrogen atoms were calculated making use of the HBUILD facility in the CHARMM program [28]. Because the reaction happens at physiologic pH, the hystidine residues in the enzyme can be protonated in several ways. The choice of the protonation state of these amino acids was based on the prior work by Diaz and Field [16]. As previously assigned by them, the residues His68, His142, His158 and His260 were protonated at nitrogen δ while His39, His62, His131 and His294 were protonated at nitrogen ε. The residue His87 was doubly protonated. To neutralize the positive net charge of the obtained structure we placed four Cl^- anions in the proximity of positive charged residues and as far as possible from the active center of PKA.

The origin of the system was then adjusted to the geometrical center of the $C_{\gamma\text{-}Asp166}$, $P_{\gamma\text{-}ATP}$, and $O_{\gamma\text{-}Ser17\_kemp}$ group of atoms (atoms in the Michaelis complex active core). The energy of the atoms within a sphere of 20 Å around this point was minimized by 50 steps. Then, a box of water molecules centered at the origin was added. The initial dimensions of the water box were 82 Å × 68 Å × 65 Å to ensure a distance of at least 10 Å between the enzyme and the box edges. All the water molecules of the box within 2.5 Å of any non-hydrogen atom of the solute (PKA, ATP, Kemptide, Mg^{++} ions), crystallographic waters, or anions were removed. The resulting system includes the monomer of PKA (342 aminoacids, AA), one ATP molecule, the Kemptide molecule (7 AA), four Cl^-, and a TIP3P water box with 10,775 water molecules, so the total number of atoms reaches 38,134.

[2] The numbers of the residues of the substrates or inhibitor are referred to the peptide bound to the enzyme in the crystallographic structures. Thus the first residue of PKI(5–24), of SP20, and of the model substrate is Thr1, and the last one is Asp20 in place of Thr5 and Asp24, respectively. In the same way the first residue of Kemptide is Leu13 and the last one is Gly19.

Simulations

The Michaelis complex represented by the model system was simulated by carrying out a classical molecular dynamics simulation with the all-atom CHARMM22 force field [29]. For the phosphorylated Thr197 we have used the standard parameters included in the CHARMM version 31 (c31) parameter files, which were optimized against experimental data and *ab initio* calculations [30].

We first removed the close contacts and the repulsive orientations of the system by carrying out energy minimizations in two steps. First, 200 steps of energy minimization of the solvent molecules (waters plus Cl^-), followed by 100 steps of optimization of all the atoms coordinates. The energy minimization method was the adopted basis set Newton–Raphson (ABNR) implemented in CHARMM [28] version c31.

Finally, starting from the resulting structure, we carried out molecular dynamics (MD) simulations with periodic boundary conditions (PBC) in the isothermal-isobaric ensemble at 296 K and 1 atm. These simulations were performed with the CRYSTAL routine of CHARMM c31: the program generates 26 images identical to the primary cell (complex plus water box) all around it; just the images within a cutoff distance (20 Å), established to decrease the number of nonbonded interactions, were built. In this work a spherical cutoff of 13 Å together with a switching function to fade the interaction energy to zero were used for the non-bonded interactions. The switching function acts between 11.5 and 12.5 Å. On the base of neutral-groups separations, the nonbonded pair list was updated every 25 steps while the image list was refreshed every 100 steps. To compute the contributions of long range electrostatic interactions, we used the Particle Mesh Ewald method with a grid spacing of $\sim$1 Å and a fourth-order B-spline interpolation to calculate the potential and forces between grid points.

The propagation of the equations of motion was accomplished with the leapfrog integration scheme, a 2 fs time step, and the extended system constant pressure and temperature algorithm implemented in CHARMM. All bond lenghts were constrained by SHAKE and the dielectric constant was set to 1.

The initial temperature was gradually raised from 0 to 296 K with 68 ps of restrained MD simulations. Harmonic potentials centered on the crystallographic positions were used for the backbone atoms of the protein and the substrate, for the heavy atoms of the ATP molecule and for the two Mg^{++} ions. Then the restraints were gradually removed

during 100 ps of simulation at 296 K, which was followed by a MD simulation of the unrestrained system for 100 ps more. The resulting structure was taken as the starting point for all the data presented in the present work. The total length of the reported free MD simulation was approximately 6 ns. Structural data was saved every 100 steps of this simulation for later analysis.

Results and discussion

In this section the results of the 6 ns MD simulation of the PKA–Mg_2ATP–Kemptide ternary complex are presented. During the simulation we have monitored several geometric parameters in order to analyze the binding of the substrate, the stability of the complex, and the interactions in the active center, as discussed in the following paragraphs.

First, the goodness of the model has been asserted by the analyses of the coordination sphere of the two Mg^{++} ions and the H-bond interactions between ATP and PKA. The average distances corresponding to the Mg-ligand bonds along the MD trajectory are given in Table S1 of the Supporting Information. Those average distance values are compared with the analogous distances in three X-ray structures. Considering that all the crystallographic structures of PKA have been obtained with two Mn^{++} ions instead of two Mg^{++} ions, all the interactions presented in Table S1 have average distances that compare quite well with the reference values. In addition, the fluctuations of those distances are lower than 5.4%, and in many cases lower than 3.5%. In Table S2 of the Supporting Information, the main H-bonds between ATP and several residues of PKA are presented. Those values present a general good agreement with the same interactions given by Diaz and Field [16] in the ternary complex with the model substrate defined in Table 1. Also, another hint of the goodness of the model comes from the comparison with the results of Cheng et al. [25], in regards to the enzyme conformation. This authors have demonstrated the dependence of the enzyme reactivity to the interaction between the phosphorylated Thr197 in the activation loop and the catalytic loop residue Arg165. In our simulation, the average distance between the phosphate oxygens of pThr197 and the guanidinium nitrogens of Arg165 are 2.72 ± 0.21 Å and 2.73 ± 0.23 Å, which together with occupancy values of ~1 indicate the stability of these important hydrogen bonds along all the simulation.

The segment from residue 301 to residue 350 constitutes the C-terminal tail of PKA. This tail is distinctly different in the various kinases and it has been proposed to selectively recognize the substrate of each of them [31]. In this region, in the case of PKA, an acidic cluster is positioned (Asp-Asp-Tyr-Glu-Glu-Glu-Glu, from residue 328 to residue 334) that, even without direct interactions with ATP and Kemptide, facilitates the binding of the P-3 and P-2 residues of the substrate and of the ATP molecule [31, 32]. An exhaustive work on mutagenesis characterization of the C-terminal tail carried out by Batkin and Shaltiel [31] shows the importance of many residues of this region: the alanine substitution of some residues may result in a complete loss of the enzymatic activity and/or in a great reduction of affinity toward the substrate and the ATP molecule. This is the case of Tyr330 that lays in the acidic cluster and whose substitution leads to both the mentioned events. As supposed by Johnson et al. [4], Tyr330 can also act to correctly orient the big and small lobes in the closing procedure during the formation of the prereactive complex. As a consequence, the distance of Tyr330 from Glu170 (one of the residues directly interacting with the P-3 arginine in the catalytic core) has been used to check the evolution of the complex conformation. The average value obtained for this distance in our simulation is 8.7 ± 0.8 Å, which is a slightly higher value than the experimental one given in the work of Johnson et al. for the closed structure. Taking into account these three factors, the presence of the Kemptide substrate in place of the PKI(5–24) inhibitor in our simulation, the different temperature (296 K instead of the crystallization temperature), and the relaxation of the complex due to the long molecular dynamics simulation, the calculated value can be considered in good agreement with the crystallographic result.

Second, the stability and the interactions of the PKA–Mg_2–ATP–Kemptide prereactive or Michaelis complex have been analyzed, which is the main objective of our long MD simulation. In Fig. 4 the root mean square deviations (RMSDs) of PKA and Kemptide backbones are depicted, using as reference structure the protein and ligand backbone coordinates in the crystallographic structure. A noticeable movement of the entire complex is observed at the beginning of the simulation while the relaxation process takes place. However after 1 ns the protein structure reaches a stable configuration. Figure 4b shows the trend of the Kemptide RMSD: it can be observed that after the initial great movement of the entire peptide, it starts to level off but, at about 2 ns, it does a sudden jump. Then it begins again to stabilize itself and finally comes, after 3 ns, to a quite steady situation. In order to explain this unusual trend of the Kemptide RMSD we have analyzed it more carefully. We have considered the entire peptide backbone and three parts of it, the N-terminus (Leu13-Arg14), the C-terminus (Leu18-Gly19) and the central part (Arg15-Ala16-Ser17). In Fig. 5a the heavy-atoms RMSDs of the Kemptide substrate structure are presented. As inferred from it, the C-terminus and the central part of the peptide present almost constant RMSDs. Thus the global behaviour

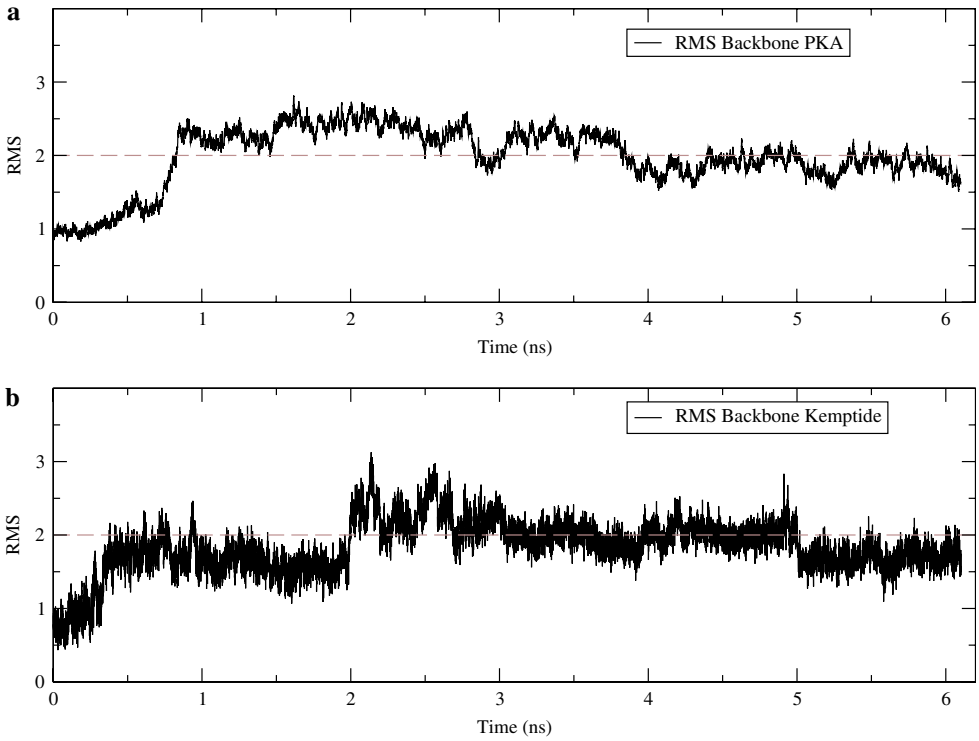

Fig. 4 PKA backbone (**a**) and Kemptide backbone (**b**) RMSDs

is essentially due to the N-terminus and particularly to the Leu13, which is quite free to move and needs a longer time to settle down in a stable conformation. The corresponding RMSDs average values calculated over the last 3 ns are: 1.9 ± 0.2 Å for the entire peptide backbone, 1.8 ± 0.5 Å for the N-terminus, 1.9 ± 0.2 Å for the C-terminus, and 1.9 ± 0.2 Å for the central part.

In Fig. 5b, the backbone of our substrate Kemptide and the backbone of the larger inhibitor molecule PKI(5–24) are presented. The figure has been obtained after the alignment of the two structures with respect to the backbone position, followed by a small rigid translation. From the Kemptide N-terminus (on the left), corresponding to the residue 13 of PKI(5–24), to the Kemptide C-terminus (on the right), corresponding to the residue 19 of PKI(5–24), the conformation of the two peptides is compared. As it can be seen, after 6 ns of MD simulation the Kemptide substrate backbone, that has already reached a stable conformation, maintains the original geometrical conformation of the pseudo-consensus sequence (Arg14-Arg15-Asn16-Ala17-Ile18) of PKI(5–24) from which it was modeled. The only remarkable difference between the small heptapeptide substrate conformation and the corresponding seven-residue structure in PKI(5–24) consists in the position of Gly19 (C-terminus in Kemptide). This is probably due to the absence of important interactions of this residue with PKA, whereas His19 of PKI(5–24) interacts with Ser53 of the enzyme. Importantly, the only common residues between Kemptide and the pseudo-consensus sequence are Arg14 and Arg15. This fact supports the apparent requirement of the two arginines as specificity determinants of PKA substrates. In fact, as indicated in the Introduction section, peptide substrates of PKA are characterized by the presence of two arginine residues in the P-2 and P-3 positions, where P is the phosphorylation site, and by the presence of a hydrophobic residue in the P+1 position [7–11]. As will be shown in the next paragraph, the guanidinium and backbone atoms of the P-3 and P-2 Arg residues and the carbonyl group of the P+1 Leu residue constitute important anchorage points between the Kemptide substrate and the PKA enzyme, and are dynamically stable interactions along the simulation. In addition, as can be seen in Fig. 6, the conformation of the side chain of the P-site Ser corresponds to the dominant G^+ rotameric state as found by Cheng et al. [25] in their simulation of the activated PKA enzyme with the SP20 substrate.

In Table 2 the average distances of the main interactions between Kemptide and PKA residues are given, along with their occupancies, lifetimes, and the events corresponding to the formation of those H-bonds. From those data it can be inferred that, in agreement with experimental observations on the crystal structure of the ternary complex PKA–Mn_2–AMP–PNP–PKI(5–24) [6], Arg14 mainly interacts with Glu127 and Thr51, whereas Arg15 interacts with Glu170 and Glu230. All these interactions are rather stable hydrogen bonds of O–N average distances between 2.73 and 3.38 Å. We have also analyzed the distances between residues Tyr204 and Glu230, because the results obtained

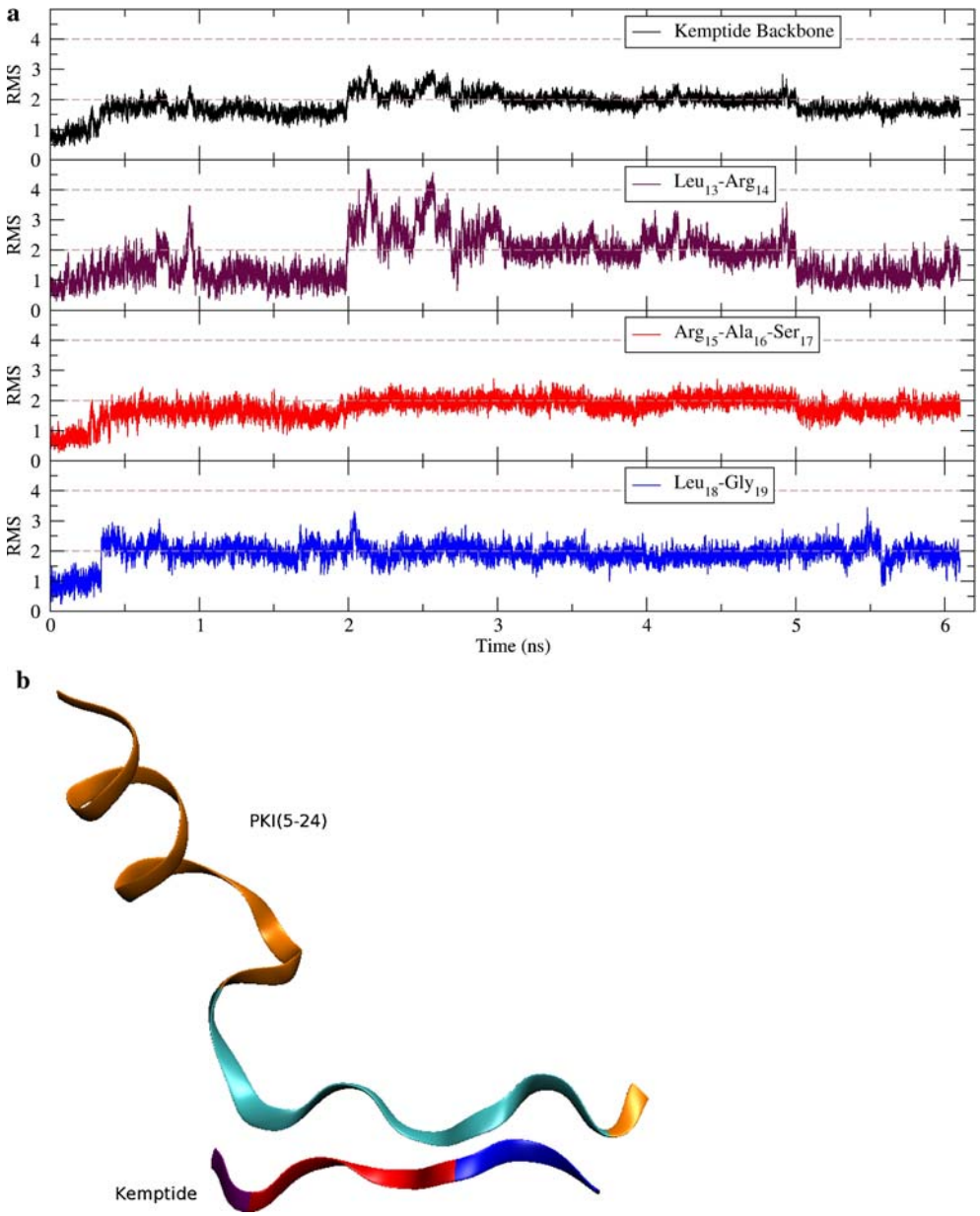

Fig. 5 (**a**) Kemptide RMSDs are shown. From top to the bottom: the complete substrate, the N-terminus Leu13-Arg14, the central residues Arg15-Ala16-Ser17, and the C-terminus Leu18-Gly19. (**b**) Graphic representation of the protein kinase inhibitor PKI(5–24) bound to the enzyme in the 1CDK crystallographic structure in comparison to Kemptide after ~6 ns of simulation. The colors of the three sections of Kemptide correspond to those of the RMSDs in (**a**)

when mutating the former residue to alanine indicate that Tyr204 contribute to peptide binding. This fact has been attributed to its hydrogen bond interaction with Glu230 [9] that allows a possible interaction between P-2 Arg and the π electrons of the aromatic ring of Tyr204. The average distance of this interaction is 3.28 ± 0.89 Å in our simulation. Only the interaction between the backbone carbonyl of Arg15 and the sidechain of Lys168 in the PKI(5–24) complex is not present in our simulation of the Kemptide–protein complex. On the other hand, the backbone amino group of Leu18 interacts with the carbonyl group of Gly200. The hydrophobic side chain of this residue partially lies in the hydrophobic environment provided by the side chains of Pro202 and Leu205.

In agreement with the results obtained by Diaz and Field [16] in the study of their 20-residue peptide, it turns out that the serine nucleophile is positioned in the PKA active site by means of ion-pairs, hydrogen bonds, and hydrophobic interactions that involve the residues of the peptide consensus sequence. Particularly, the two Arg residues and the Leu residue are involved. However, Kemptide is not so firmly bound as the model peptide used

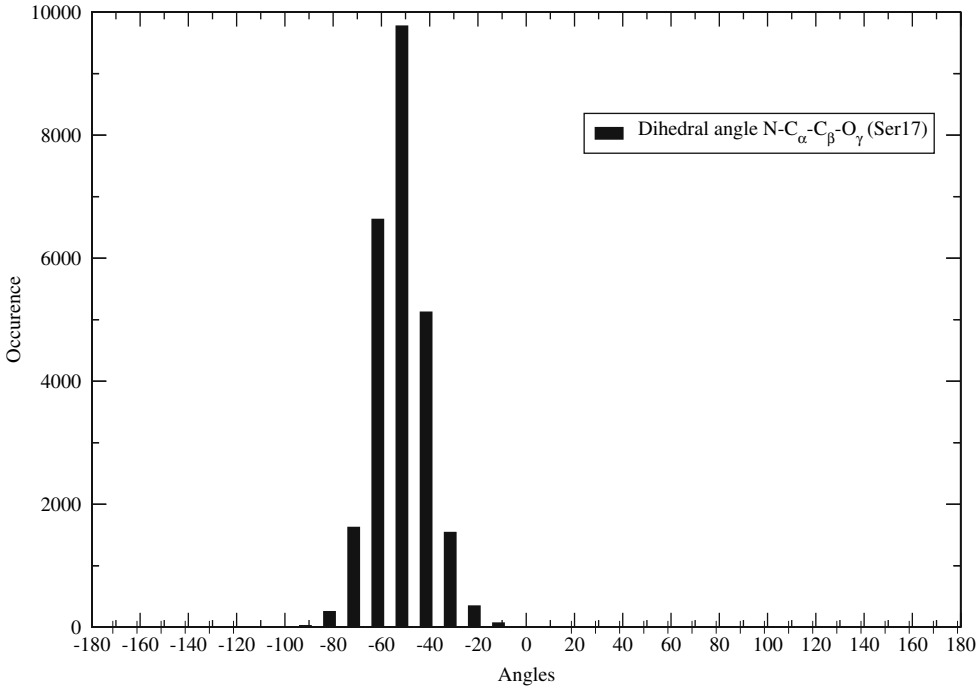

Fig. 6 The side-chain torsion angle distribution of the P-site Ser17

by Diaz and Field [16] to the PKA active site (compare hydrogen-bond distances and occupancies in Table 2, and Table S4 of that reference [16]); this is probably due to the absence in our substrate of the helix that, in PKI(5–24), helps the peptide to tight bind the enzyme and so to act as an inhibitor.

Third, our results can also give some insights into the nature of the reaction mechanism catalyzed by PKA. As already indicated, nucleophilic substitutions at phosphorus have been described in terms of a continuum of mechanisms with two limiting cases, a fully dissociative mechanism that proceeds via a trigonal metaphosphate

Table 2 Average distances with their fluctuations, occupancies, lifetimes, events of H-bonds and reference distances of all the interactions observed between PKA and the substrate

Pair	Distance, Å	Occ.	Lifetime, ps	Events	Ref., Å
$O_\varepsilon 2(GLU_{170})$–$N_\varepsilon(ARG_{15})$	2.90 ± 0.18	0.98	9.62	626	2.89
$O_\varepsilon 2(GLU_{170})$–$N(ARG_{15})$	3.38 ± 0.56	0.52	3.97	799	2.66
$O_\varepsilon 1(GLU_{170})$–$N_\eta 2(ARG_{15})$	2.74 ± 0.12	0.97	68.41	88	2.96
$O_\varepsilon 1(GLU_{127})$–$N_\varepsilon(ARG_{14})$	2.82 ± 0.16	0.98	23.42	258	3.05
$O_\varepsilon 2(GLU_{127})$–$N_\eta 2(ARG_{14})$	2.73 ± 0.11	0.77	36.71	130	2.89
$O_\gamma(SER_{53})$–$O(SER_{17})$	3.98 ± 0.47	0.09	9.19	58	2.75
$O(GLY_{200})$–$N(LEU_{18})$	2.90 ± 0.19	0.96	10.41	567	2.72
$O(THR_{51})$–$N_\eta 1(ARG_{14})$	2.88 ± 0.18	0.94	3.96	1419	2.86
$O_\varepsilon 1(GLU_{230})$–$N_\eta 2(ARG_{15})$	3.28 ± 0.40	0.44	1.33	1972	3.31
$O_\varepsilon 1(GLU_{230})$–$N_\eta 1(ARG_{15})$	3.02 ± 0.36	0.68	1.98	2068	2.87
$O(THR_{51})$–$N_\eta 2(ARG_{14})$	3.14 ± 0.30	0.64	0.86	4607	2.90
$O_\delta 2(ASP_{166})$–$O_\gamma(SER_{17})$	3.99 ± 0.58	0.10	2.83	223	2.46
$N_\zeta(LYS_{168})$–$O_\gamma(SER_{17})$	2.88 ± 0.15	0.92	5.29	1073	3.51
$O_\gamma 1(THR_{201})$–$O_\gamma(SER_{17})$	3.03 ± 0.40	0.02	0.37	254	3.47
$O_\delta 1(ASP_{166})$–$O_\gamma(SER_{17})$	4.13 ± 0.43	0.00	0.22	74	3.49

All the distances are between heavy atoms. Occupancy (Occ.) is the percentage of the time passed at interaction distance over the total time (the cutoff distance has been chosen as 2.4 Å). Lifetime is the average time passed at bonding distance between two subsequent separations. Events is the number of times the atoms form a H-bond. The reference distances are taken from the initial crystallographic structure. The total evolution time is ~6 ns

transition state and a fully associative mechanism that involves a trigonal bipyramidal transition state. A useful distance obtainable from X-ray or NMR data, which provides a lower limit to the fractional associativity of a mechanism, is the reaction coordinate distance, defined as the distance between the entering atom and the phosphorous undergoing substitution before the reaction has begun.

Again, assuming that the only motion during the reaction is that of the [PO_3] moiety, a quantitative model for the reaction coordinate distances based on simple geometric principles has been proposed by Mildvan [33]. The Mildvan concept is that, in a dissociative transition state mechanism, the reaction coordinate distance should be greater than or equal to 4.9 Å. For a transition state with great associative character, the reaction coordinate distance could be less than 3.3 Å (van der Waals contact distance), indicating compression. High resolution X-ray and NMR structures of various protein kinase complexes with nucleotide and peptide substrate analogs have yielded conflicting values for such distances, although these studies have, by necessity, employed unreactive substrate analogs that could have affected the results. In the already mentioned crystal structure of the ternary complex PKA–Mn_2–AMP–PNP–PKI, the distance from the oxygen of the modeled serine hydroxyl group to the γ-phosphorous atom is 2.7 Å [6], in agreement with the reaction coordinate distance obtained in another complex where the C-subunit of PKA was co-crystallized with a 20-residue substrate (SP20) and MgADP [8]. Although those reaction-coordinate distance values measured from static molecular modeling are too short in accordance with Diaz and Field [16], also in a recent crystal structure of a transition state mimic of the catalytic subunit of PKA, in complex with Mg_2ADP, aluminium fluoride, and SP20, an Al–O distance of 2.3 Å was observed with the aluminium atom showing a bipyramidal coordination [1]. In the same study, the authors indicate that one of the most surprising findings is that PKI(5–24) closely resembles the transition state analog when it is bound to the catalytic subunit in a ternary complex with Mg_2ATP. Much longer reaction coordinate distances have been estimated by NMR measurements and from crystal structures of other kinases. In particular, an NMR study by Granot et al. [34] gave a distance between the γ-phosphorous and the serine hydroxyl oxygen of 5.3 ± 0.7 Å, which is in a range that would permit a dissociative mechanism. In any case, the same authors indicate that an associative mechanism remains possible because the intersubstrate distances may decrease as the transition state is approached.

Theoretical calculations have also given contradictory results concerning the reaction coordinate distance and the degree of associativity of the phosphorylation mechanisms. However, three of the most complete theoretical studies recently published agree to propose a mostly dissociative mechanism with average reaction coordinate distances which turn out to be 3.8 ± 0.3 Å [16] and 3.72 ± 0.42 Å [19, 25] in their works. The calculated average reaction coordinate distance ($P_{\gamma\text{-}ATP}$–$O_{\gamma\text{-}Ser17}$) from our simulation of the PKA–Mg_2–ATP–Kemptide prereactive complex is 4.21 ± 0.58 Å, in agreement with the previously mentioned theoretical works on a different substrate complex. In Table 3 the occupancy, lifetimes, and events corresponding to this interaction are also shown, along with the same parameters for the $O3_{\gamma\text{-}ATP}$–$N_{\gamma\text{-}Ser17}$ and the $O1_{\gamma\text{-}ATP}$–$O_{\gamma\text{-}Ser17}$ interactions. In Fig. 7 the trend of those three distances along the 6 ns simulation is depicted, together with the corresponding trends of other interactions. Importantly, the evolution of the reaction coordinate distance values suggests that the system oscillates between configurations compatible with both the associative and dissociative mechanisms, which explains the intermediate value of the average distance obtained in our simulation. In summary, these results, in addition with the rest of PKA–Kemptide interactions, indicate that Kemptide is loosely bound in the active site of PKA, which is in good agreement with the $K_m$values experimentally measured.

Finally, we have centered our attention on the interactions of Ser17, Asp166, and Lys168 attempting to elucidate which are the most important actors in the phosphorylation process.

Asp166 is a conserved residue in the active site of all protein kinases and resides in a position likely to form hydrogen bonds with the substrate serine. Mutation to alanine causes at least a 300-fold reduction of v_{max} without greatly affecting the K_m for either ATP or the substrate peptide. It was first suggested to be the catalytic base, which accepts the proton from the substrate hydroxyl and enhances the nucleophilic reactivity of the P-site serine [17]. However, considering the acidities of aspartic acid ($pK_a = 3.80$) and serine ($pK_a = 14$) in the aqueous solvent, it was also suggested that Asp166 might not be an effective general base and that its unique catalytic role might be related with substrate orientation [17]. As indicated in the Introduction section, the theoretical results of Diaz and Field [16] and Cheng et al. [19, 25] corroborate the role as catalytic base for Asp166 but with the proton transfer taking place in a late stage of the phosphorylation process. In a study of the molecular electrostatic properties of PKA residues, Blachut-Okrasiska et al. [21] found an exaggerated shift to negative values for the Asp166 pK_a. This was interpreted as an indication that the group remains fixed in a single ionization state at any pH for which the complex is actually stable in solution.

In our simulation of the PKA–Mg_2ATP–Kemptide prereactive complex the $O_{\delta}1_{Asp166}$–$O_{\gamma\text{-}Ser17}$ and $O_{\delta}2_{Asp166}$–$O_{\gamma\text{-}Ser17}$ average distances are 4.13 ± 0.43 Å and 3.99 ± 0.58 Å, respectively (see Table 2 and Fig. 7).

Table 3 Average distances with their fluctuations, occupancies, lifetimes, events of H-bonds and reference distances of all the interactions observed between ATP and the substrate

Pair	Distance, Å	Occ.	Lifetime, ps	Events	Ref., Å
$P_{\gamma}(ATP)$–$O_{\gamma}(SER_{17})$	4.21 ± 0.58	–	–	–	3.22
$O1_{\gamma}(ATP)$–$O_{\gamma}(SER_{17})$	4.50 ± 0.68	0.04	1.37	177	3.00
$O3_{\gamma}(ATP)$–$N(SER_{17})$	4.63 ± 0.46	0.04	1.36	173	3.02

All the distances are between heavy atoms. Occupancy (Occ.) is the percentage of the time passed at interaction distance over the total time (the cutoff distance has been chosen as 2.4 Å). Lifetime is the average time passed at bonding distance between two subsequent separations. Events is the number of times the atoms form a H-bond. The reference distances are taken from the initial crystallographic structure. The total evolution time is ∼6 ns

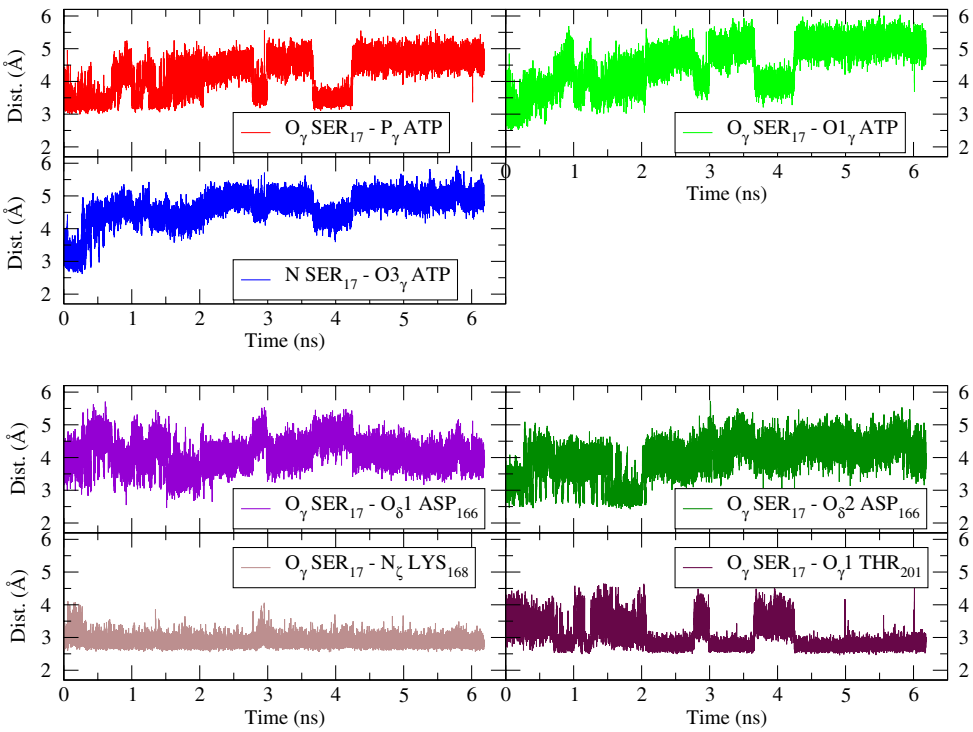

Fig. 7 Evolution of the interaction distances involving Ser17 through the MD simulation. In the upper part the interaction distances of Ser17 with ATP are represented. In the lower part the interactions of Ser17 with some residues of PKA are depicted

These long distances indicate a weak interaction between Asp166 and Ser17, but the great fluctuations of those distances (more than 10% of the average value) do not exclude the possibility of a stochastic closer contact. Importantly, the shorter value obtained in our 6 ns simulation of the PKA–Mg_2ATP–Kemptide complex is in good agreement with the average distance (3.77 ± 0.61 Å) obtained for the G^+ rotamer of Ser17 in the study of Cheng et al. [25], where they have carried out a 25 ns MD simulation with the SP20 substrate bound. In Table S3 other interactions between Asp166 and several PKA residues in the active center are analyzed. Among those interactions, the H-bonds between Asp166 and Lys168 are the most stable along the simulation process. The $O_{\delta}1_{Asp166}$–N_{Lys168} and $O_{\delta}2_{Asp166}$–N_{Lys168} average distances are the shortest ones, with values of 2.86 ± 0.22 Å and 2.85 ± 0.17 Å, respectively, and present the highest occupancies (0.91 and 0.74, respectively).

The other residue that has been proposed to be important in the phosphorylation reaction catalyzed by PKA is Lys168. This is a highly conserved residue, present in all the serine and threonine kinases, but not in tyrosine kinases. Almost along the entire simulation of the prereactive ternary complex, Lys168 lingers on maintaining a stable interaction not only with Asp166 (see Table S3), but also with Ser17 (see Table 2) and with one of the oxygens of the γ-phosphoryl group (see Table S2, Fig. S2). These three interactions present occupancies that are higher than 90% and really small distance fluctuations during the equilibration. The average distances $N_{\zeta\text{-}Lys168}$–$O_{\gamma\text{-}Ser17}$ and $O2_{\gamma\text{-}ATP}$–$N_{\zeta\text{-}Lys168}$ are 2.88 ± 0.15 Å and 2.81 ± 0.16 Å, respectively.

Due to the position of Lys168 in the active core and to the stable interactions with the surrounding molecules (ATP, Ser17 and Asp166) we have considered a reaction pathway different from the previously proposed ones: the

similarity between the catalytic triad of the serine protease chymotrypsin and the three residues just named makes us suppose a similar mechanism in which the lysine acts as a bridge between Asp166 and the serine residue. Due to its interaction with the negative charged side chain of the aspartate, Lys168 can move its protons passing one of them to Asp166 and recovering another one from Ser17. Thus Ser17 would become highly reactive as a nucleophile and attack the γ-phosphoryl group of ATP. At the moment this is just a hypothesis and we should carry out a complete reaction dynamics study to say something more about it. Similarly, it has been supposed by other research groups that Lys168 can give one of its proton to the oxygen of ATP and then recover another one from Ser17 [19].

Conclusions

In the present work we have simulated the complex between the synthetic hepta-peptide Kemptide, protein kinase A, and Mg_2ATP, in order to build a model comparable to the experimental kinetic and thermodynamic data found in the literature for this enzyme, which is a prototype of the kinase super-family. In fact, there are no crystall structures of PKA kinase with such a small peptide.

Starting from the PDB structure 1CDK, we have built up our system by substituting the Mn_2AMP-PNP molecule with Mg_2ATP and modeling the synthetic hepta-peptide Kemptide on the backbone of the PKI(5–24) inhibitor bound to the crystallographic structure.

One of the main targets of this work is to attest the stability of the manually set up system, according with the crystallographic structure and with previous works of other research groups on models with different substrates. The results obtained from the RMSD analyses of Kemptide and PKA backbones, and from the interaction distance analysis of several active site residues indicate that the simulated substrate–enzyme complex mimics the substrate binding pattern known for PKA.

The analysis of prereactive distances does not clarify which of the two proposed pathways (associative and dissociative) for the enzymatic phosphorylation is the actual one: the large Ser17-Asp166 distance suggests a really feeble interaction between the residues, but the big fluctuations of this distance do not definitively exclude the dissociative mechanism (Scheme B in Fig. 2). On the other hand, the Ser17-ATP prereactive distance trend along the simulation indicates that the system oscillates between configurations compatible with each of the proposed mechanisms, according to the value of the reaction coordinate that is proposed to characterize each of them in previous experimental works.

Finally, due to the particular conformation of the active site, we suggest an alternative mechanism analogous to that of the serine proteases. The short and stable distances between Lys168 and Ser17, ATP, and Asp166 seem to indicate a major role of the lysine residue in the reaction. We think that Lys168 can act as a bridge between Asp166 and Ser17. It can pass a proton to the aspartate residue and then extract the alcoholic proton from the serine residue, making it capable to nucleophilically attack the γ-phosphoryl group of ATP.

Acknowledgments We are grateful for financial support from the Spanish "Ministerio de Educación y Ciencia" and the "Fondo Europeo de Desarrollo Regional" through projects CTQ2005-07115/BQU and BIO2004-05879-C02-01, the "Generalitat de Catalunya" (2005SGR00400), and the "Universitat Autònoma de Barcelona" (EME2006-18). M.G.-V. thanks the "Ramon y Cajal" program for financial support.

References

1. Madhusudan, Akamine P, Xuong NH, Taylor SS (2002) Nat Struct Biol 9:273
2. Adams JA (2001) Chem Rev 101:2271
3. Cole PA, Courtney AD, Shen K, Zhang ZS, Qiao YF, Lu W, Williams DM (2003) Acc Chem Res 36:444
4. Johnson DA, Akamine P, Radzio-Andzelm E, Madhusudan, Taylor SS (2001) Chem Rev 101:2243
5. Gibbs CS, Zoller MJ (1991) J Biol Chem 266:8923
6. Bossemeyer D, Engh RA, Kinzel V, Ponstingl H, Huber R (1993) EMBO J 12:849
7. Gibbs CS, Zoller MJ (1991) Biochemistry 30:5329
8. Madhusudan, Trafny EA, Xuong NH, Adams JA, Teneyck LF, Taylor SS, Sowadski JM (1994) Protein Sci 3:176
9. Moore MJ, Adams JA, Taylor SS (2003) J Biol Chem 278:10613
10. Shabb JB (2001) Chem Rev 101:2381
11. Brinkworth RI, Horne J, Kobe B (2002) J Mol Recognit 15:104
12. Kemp BE, Graves DJ, Benjamini E, Krebs EG (1977) J Biol Chem 252:4888
13. Qamar R, Yoon M-Y, Cook PF (1992) Biochemistry 31:9986
14. Grant BD, Adams JA (1996) Biochemistry 35:2022
15. Kemp BE, Benjamini E, Krebs EG (1976) Proc Nat Acad Sci 73:1038
16. Diaz N, Field MJ (2004) J Am Chem Soc 126:529
17. Zhou J, Adams JA (1997) Biochemistry 36:2977
18. Williams DM, Cole PA (2002) J Am Chem Soc 124:5956
19. Cheng YH, Zhang YK, McCammon JA (2005) J Am Chem Soc 127:1553
20. Hart JC, Sheppard DW, Hillier IH, Burton NA (1999) Chem Comm 79
21. Blachut-Okrasinska E, Lesyng B, Briggs JM, McCammon JA, Antosiewicz JM (1999) Eur Biophys J 28:457
22. Sheppard DW, Burton N, Hillier IH (2000) J Mol Struct 506:35
23. Hutter MC, Helms V (2003) Int J Quant Chem 95:479
24. Valiev M, Kawai R, Adams JA, Weare JH (2003) J Am Chem Soc 125:9926
25. Cheng YH, Zhang YK, McCammon JA (2006) Protein Sci 15:672
26. Klahn M, Rosta E, Warshel A, Afklaehn M, Rosta E, Warshel A (2006) J Am Chem Soc 128:15310
27. Vanpatten SM, Ng DC, Thng JPH, Angelos KL, Smith AJ, Walsh DA (1991) Proc Nat Acad Sci 88:5383

28. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comput Chem 4:187
29. MacKerell AD Jr, Bashford D, Bellott M Jr, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) J Phys Chem B 102:3586
30. Feng MH, Philippopoulos M, MacKerell AD, Lim C (1996) J Am Chem Soc 118:11265
31. Batkin M, Shaltiel ISS (2000) Biochemistry 39:5366
32. Shaltiel S, Cox S, Taylor SS (1998) Proc Nat Acad Sci 95:484
33. Mildvan AS (1997) Proteins 29:401
34. Granot J, Mildvan AS, Bramson HN, Kaiser ET (1980) Biochemistry 19:3537