Role of phosphorylated Thr160 for the activation of the CDK2/Cyclin A complex

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · NOVEMBER 2005

Impact Factor: 2.63 · DOI: 10.1002/prot.20697 · Source: PubMed

5 AUTHORS, INCLUDING:

Marco De Vivo
Istituto Italiano di Tecnologia
43 PUBLICATIONS **711** CITATIONS
SEE PROFILE

Andrea Cavalli
University of Bologna
171 PUBLICATIONS **4,896** CITATIONS
SEE PROFILE

Giovanni Bottegoni
Istituto Italiano di Tecnologia
45 PUBLICATIONS **795** CITATIONS
SEE PROFILE

Paolo Carloni
Forschungszentrum Jülich
276 PUBLICATIONS **5,974** CITATIONS
SEE PROFILE

Role of Phosphorylated Thr160 for the Activation of the *CDK2/Cyclin A* Complex

Marco De Vivo,[1] Andrea Cavalli,[1] Giovanni Bottegoni,[1] Paolo Carloni,[2]* and Maurizio Recanatini[1]*
[1]*Department of Pharmaceutical Sciences, University of Bologna, Bologna, Italy*
[2]*International School for Advanced Studies SISSA/ISAS,Trieste, Italy*

ABSTRACT **The enzymatic activity of the CDK2/Cyclin A complex increases upon the specific phosphorylation of Thr160@CDK2. In the present study, we have performed a comparative molecular dynamics (MD) study of models of the complex CDK2/Cyclin A/Substrate, which differ for the presence or absence of the phosphate group bound to Thr160. The models are based on two X-ray structures available for CDK2/CyclinA and pCDK2/CyclinA/Substrate complexes. In this way, we analyze the influence of the phosphorylated Thr160 (pThr160) on both the flexibility of CDK2 activation loop (AL) and substrate binding in CDK2. Our calculations point to a decreased flexibility of the AL in the phosphorylated model, in fairly good agreement with experimental data, and to a key role of pThr160 for substrate recognition and stability. Multiple alignments of the CDKs sequences point to the very high conservation of the AL sequence among the CDKs, thus extending our results to all CDKs. Proteins 2006;62:89–98.** © 2005 Wiley-Liss, Inc.

Key words: cyclin-dependent kinase; CDK; CDK2/cyclin A complex; Thr160; phosphorylation; molecular dynamics

INTRODUCTION

Cyclin-dependent kinases (CDKs) assist the γ-phosphate transfer from ATP to peptide substrates belonging to downstream targets such as the tumor-suppressor proteins, pRb and the related p107 and p130, and E2F transcription factor.[1,2]

The CDK phosphorylation process controls signal transduction pathways essential for the initiation, progression, and completion of eukaryotic cell cycle.[3] Aberrant activities of CDKs have been found in a number of diseases including neurodegenerative disorders (e.g., Alzheimer's disease, lateral sclerosis, and stroke), cardiovascular disorders (e.g., atherosclerosis), viral infection, cancer, and alopecia.[4]

CDKs require a two-step process to be fully functional: (1) The association with the regulatory cyclin subunit; (2) the phosphorylation of a specific residue (Thr), situated along the so-called activation loop (AL) of CDK, by the CDK-activating kinase protein (CAK) (Fig. 1).[5–7] Besides, other regulatory pathways are used to coordinate CDK-mediated (anti)proliferative signals: regulated assembly of holoenzymes, tight-binding inhibitors, ubiquitin-mediated proteolysis of Cyclins, and CDK native inhibitors (e.g., Cip and INK4 inhibitors families).[1,8]

To date, the structures of three members (CDK2, CDK5, and CDK6) have been determined by X-ray diffraction. We focus here on CDK2, since it is the most extensively studied both structurally and biochemically.[9] The activity of this protein alone is negligible, while the CDK2/Cyclin A binary complex and the pCDK2 alone exhibit about 0.2 and 0.3%, respectively, of the activity of the fully activated pCDK2/Cyclin A binary complex.[10]

The fold of CDK2 (Fig. 2) consists of: (1) the N-term domain (residues 1–82), mostly composed by β-sheets and one α-helix with the "PSTAIRE" sequence (residues 46–56), which is a fingerprint of this class of proteins[11] and constitutes the main point of interaction with Cyclin A; (2) the hinge region (residues 83–87) connecting the N-term to the C-term; and (3) the C-term domain (residues 88–298) formed principally by α-helices and by the AL (residues 145–172), which contains the specific residue (Thr160) that is being phosphorylated by CAK for activation. A number of loops are present in the entire CDK2 structures, of which we consider the eight most relevant in our study (L1-L8, Fig. 2).

The determination of the structures of CDK2 in the free state[9] and in complex with Cyclin A,[11] either possessing phosphorylated Thr residue (pThr160) or unphosphorylated group[10] (see Table I of the Supplementary Material), has allowed us to propose a molecular mechanism for the process leading to protein activation (Steps 1–2 above): Cyclin A binding causes a rotation of the PSTAIRE helix, which brings Arg50 of the fingerprint region close to the twice negatively charged pThr160. Because of the Arg-pThr160 electrostatic interaction, also pThr160 moves from its position along residues interacting with it:

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585.suppmat/

Grant sponsor: INFM; Grant sponsor: MIUR-COFIN.

Marco De Vivo's current address is Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104-6323.

*Correspondence to: Prof. Paolo Carloni, International School of Advanced Studies, SISSA/ISAS, Via Beirut 2-4, 34014 Trieste, Italy. E-mail: carloni@sissa.it and Prof. Maurizio Recanatini, Department of Pharmaceutical Sciences–University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy. E-mail: maurizio.recanatini@unibo.it

Received 17 May 2004; Revised 12 May 2005; Accepted 2 June 2005

Published online 16 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20697

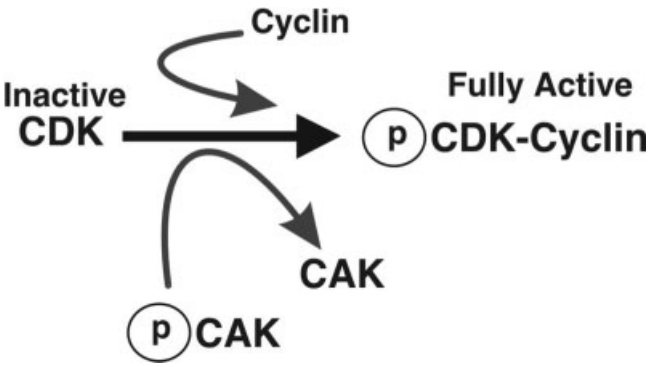


Fig. 1. CDK regulatory mechanism. Schematic illustration of the major CDK regulatory mechanism, which is composed by a two-step process: the Cyclin binding and the phosphorylation along the AL by the CDK-activating kinase protein (CAK).

Arg150@CDK2 (from the AL), Arg126@CDK2 (from the catalytic site), and a basic residue (Lys or Arg) of the substrate peptide (P+3@Sub), located three residues after the phosphorylable serine. Thus, the motion of pThr160 causes a large rearrangement of the AL from a closed to an opened conformation in which AL does not block the catalytic cleft. This enables the catalytic site region to adopt both ATP and the polypeptide substrate.

In this report, we address the role of AL flexibility[12–15] for Cyclin A binding[8,11] and substrate binding/stability[16,17] using molecular simulations. We perform a comparative molecular dynamics (MD) study on systems that differ only for the presence or absence of the phosphate group bound to Thr160. We then extend our findings by aligning two sets of sequences constituted by either the sequences considered as the phylogenetic evolution of CDKs,[18] or by the sequences of the 11 CDKs (CDK1-CDK11) forming the CDK protein family. The final aim is a deeper understanding at a molecular level of the mechanisms governing CDK's activation: this is helpful, along with information on the reaction mechanism,[19] to design therapeutically useful compounds affecting CDK's function.[20,21]

The phosphorylated system (pCDK2 hereafter) is based on the X-ray structure of phosphorylated CDK2 in complex with Cyclin A, and the HHA**S**PRK substrate analog, where **S** is the phosphorylable serine.[16] pCDK2 includes also the ATP molecule and one Mg^{2+} ion in the catalytic site.

Two models of the unphosphorylated complex are constructed either from pCDK2 (by replacing pThr160 with Thr160, CDK2 hereafter, Fig. 3) or from the X-ray structure of the unphosphorylated complex CDK2/Cyclin A[11] (CDK2-II hereafter). The latter differs from CDK2 for the absence of substrate and the catalytically essential Mg^{2+} ion.

The comparison among our simulations points to a lower flexibility of the AL in pCDK2, mostly because of the presence of electrostatic and H-bond interactions between pThr160 and Arg50, Arg126, and Arg150. These interactions are lost in CDK2, with the consequent increased mobility of the AL, as observed experimentally. The phosphorylation of Thr160 clearly plays a role also in substrate recognition and stability, mainly because of the interaction

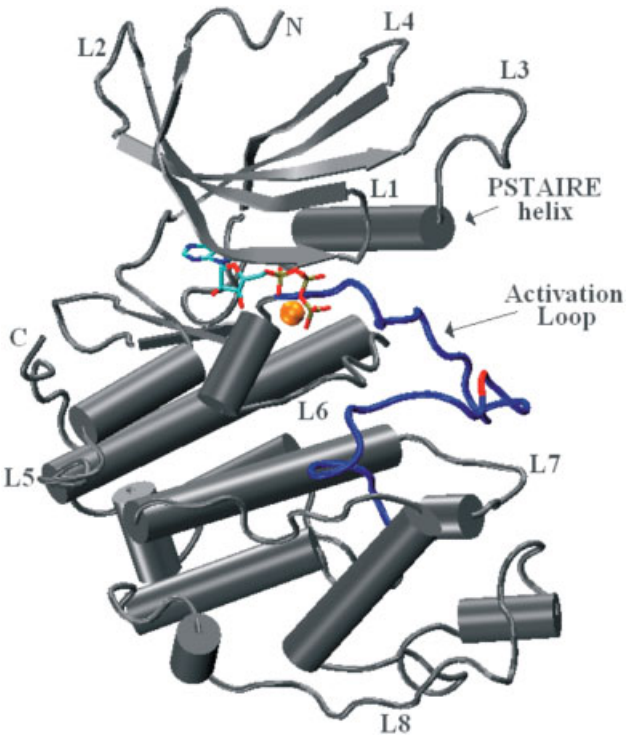


Fig. 2. CDK2 folding and binding site. Cartoons of secondary structures of CDK2 alone (PDB code 1QMZ).[16] Eight structurally relevant loops are numbered (L1–L8). Position of pThr160 is indicated in red within the AL colored in blue. ATP is represented in sticks colored based on atom-type and an orange sphere indicates the Mg^{2+} cation present in the active site. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

TABLE I. RMSD Values of the Backbone (All) Atoms of the Three Systems

System	RMSD (Å)
pCDK2	1.6 ± 0.1 (2.2 ± 0.1)
CDK2	1.8 ± 0.2 (2.5 ± 0.2)
CDK2-II	1.6 ± 0.1 (2.4 ± 0.1)

between the phosphate group and the conserved basic P+3@Sub residue (lysine), as previously proposed by Brown et al.[16] In fact, this salt bridge ensures an overall higher stability of the N-terminal part of the peptide substrate. The simulations also support an important role of Cyclin A domain in stabilizing the Thr160 in the correct conformation for its eventual phosphorylation. Moreover, the Cyclin A binding keeps AL in the opened conformation, explaining the basal enzymatic activity experimentally found for the partially activated CDK2/Cyclin A binary complex.[10] Finally, multiple sequence alignments underline a very high conservation of the specific AL sequence in both sets (see Fig. 9), suggesting a possible extension of our findings to the CDK family.

COMPUTATIONAL METHODS

Molecular Dynamics Simulations

The structure of pCDK2 has been solved at 2.2 Å resolution,[16] (PDB entry: 1QMZ). The structure contains

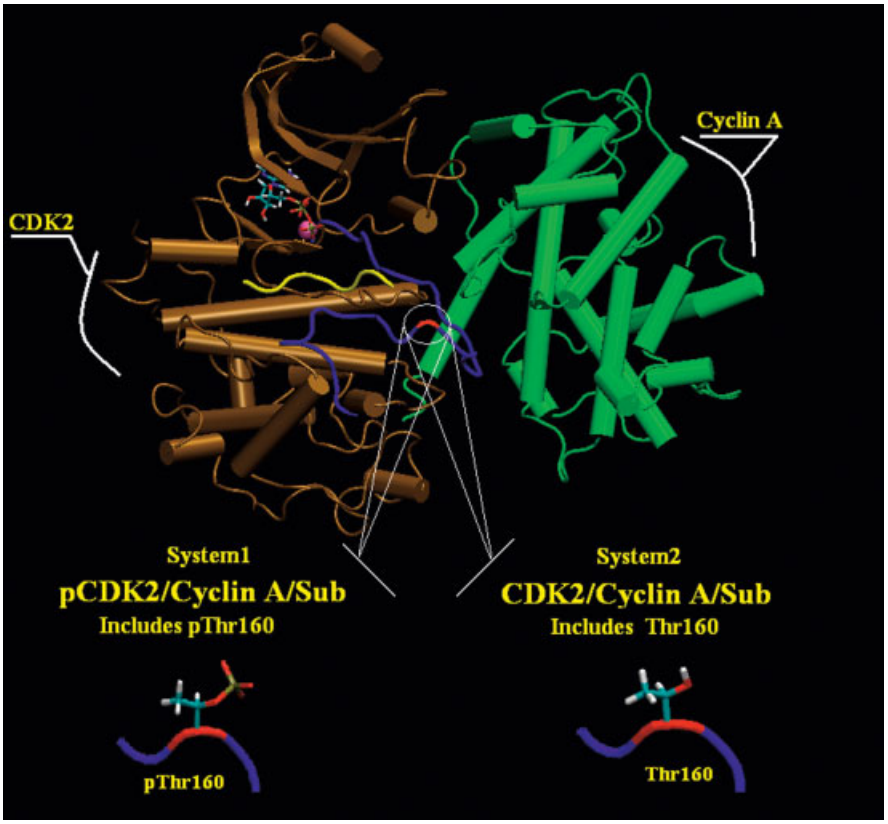


Fig. 3. Model systems. pCDK2 and CDK2 systems, are both composed of (p)CDK2, Cyclin A, and the substrate analog (Sub, in yellow). The fully active pCDK2 includes also the phosphorylated Thr160 (pThr160). In contrast, the CDK2 includes a normal Thr residue in position 160 (in red within the blue AL).

TABLE II. RMSD Values of the Backbone (All) Atoms of Protein Groups in the Three Systems

	pCDK2 system	CDK2 system	CDK2-II system
pCDK2 only	1.5 ± 0.1 (2.3 ± 0.1)	1.8 ± 0.1 (2.5 ± 0.1)	1.5 ± 0.1 (2.4 ± 0.1)
Cyclin A	1.2 ± 0.1 (1.9 ± 0.1)	1.1 ± 0.1 (1.9 ± 0.1)	1.3 ± 0.1 (2.1 ± 0.1)
PSTAIRE helix	0.5 ± 0.1 (1.5 ± 0.1)	0.4 ± 0.1 (1.3 ± 0.1)	0.4 ± 0.1 (1.3 ± 0.1)
ATP	0.7 ± 0.1	0.7 ± 0.1	1.9 ± 0.2

the pCDK2 protein (297 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), HHA**S**PRK substrate where **S** is the serine that is phosphorylated by ATP during the enzymatic catalysis. Asp and Glu groups were assumed to be ionized. Hydrogen atoms were added assuming standard bond lengths and angles. The −2 charge of pCDK2 was neutralized by adding two sodium ions using the electrostatic map potential included in xLEaP of the AMBER package.[22] The 382 crystallographic water molecules were retained.

The structure of CDK2 was constructed from pCDK2 by replacing pThr160 with Thr160. Both systems were immersed in a 82.1 × 89.2 × 100.9 Å box containing about 17,000 water molecules. The total number of atoms was about 62,000.

The structure of CDK2-II has been solved at 2.3 Å resolution,[11] (PDB entry: 1FIN). It contains the entire CDK2 (298 residues), residues 173–432 of Cyclin A, and the ATP cofactor. The −4 charge of CDK2-II was neutralized by adding four sodium ions. Two hundred and forty-two crystallographic water molecules of the crystal structure complex were retained. This system was immersed in a 104.7 × 90.6 × 81.1 Å box containing about 18,000 water molecules. The total number of atoms was about 63,500.

The calculations were performed using the SANDER program included in the Amber7 package.[22] The AMBER force field[23] for the proteins, substrate, and counterions was used. Ad hoc charges for the CDK's catalytic site including ATP, Mg^{2+}, Asp145, Asn132, and a water molecule, reported in Cavalli et al.,[24] were implemented. Also, ad hoc charges for pThr160 of pCDK2 were calculated following the RESP[25] fitting procedure (see Table II of the Supplementary Material). The bond stretching, bending, torsional parameters, and the van der Waals parameters for all these groups except Mg^{2+} are those of the AMBER

force-field.[23] As far as the hexa-coordination shell of the Mg^{2+} cation is concerned for pCDK2 and CDK2, we followed the procedure of Cavalli et al. and Banci et al.,[24,26] by imposing harmonic restraints on the six distances constituting the octahedral geometry centered on the Mg^{2+} cation (see Fig. 1 of the Supplementary Material).

The time step for the integration of the equation of motion was set equal to 1.5 fs. Bonds involving hydrogen atoms were constrained to their equilibrium position with the SHAKE algorithm.[27] Periodic boundary conditions were applied. Long-range electrostatic interactions were calculated with the Particle-Mesh Ewald (PME) summation,[28,29] with ~1 Å charge grid spacing interpolated by fourth-order B-spline and by setting the direct sum tolerance to 10^{-5}. A cut-off equal to 12 Å was used for short-range electrostatics and van der Waals interactions. The systems were coupled to a Berendsen bath[30] in order to perform simulations at constant temperature and pressure (300 K and 1 atm, NPT ensemble). In order to minimize the energy and remove bad contacts in the initial geometry, 3,000 steps of steepest descent, followed by 4,000 steps of conjugate gradient were performed on the whole systems. Then, the systems were gently heated up to 300 K by means of 150 ps of simulations (i.e., 6 steps of 50 K of 25 ps each one). Subsequently, ~10.0 ns of classical dynamics of every system was collected.

The following properties were calculated:

1. Root mean square deviations (RMSD) of the atoms at time t with respect to the initial minimized structure:

$$RMSD(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\Delta r_i(t)^2}$$

where $\Delta r_i(t) = r_i(t) - r_i(t_0)$ is the displacement of atom i during time t and N is the number of atoms over which the average is calculated.

2. The Debye-Waller factors (B-factors) of CDK2, pCDK2, and Sub were calculated over the last 7 ns of trajectories, as:

$$\langle \Delta r_i^2 \rangle = \frac{3B_i}{8\pi^2}$$

where

$$\langle \Delta r_i^2 \rangle$$

is the mean square fluctuation of the Cα atom of residue i.

Calculated B-factors of pCDK2 and Sub were compared with those experimentally derived from the crystal structure determined by Brown et al.;[16] [see Fig. 5(A,D)]. Then, experimentally derived B-factors of unphosphorylated CDK2 of the crystal structure determined by Jeffrey et al.[11] were used for a comparison with those calculated for CDK2 in CDK2 and CDK2-II [see Fig. 5(B)].

3. The large-scale motion of the pCDK2 was investigated in terms of cross-correlated matrix. To this aim, a recently reported course grain method[31] was applied (see Fig. 7). Shortly, the deviations of the amino acids from their reference native positions can be predicted by using an extension of the Gaussian Network approach. Besides, this method introduces a novel description of the amino acids that considers the presence of effective Cβ centroids whose degrees of freedom are entirely controlled by the Cα atoms. This model showed its powerful as a simple tool for a predictive characterization of proteins near-native motions.

TABLE III. H-Bond Network Shown in Figure 6†

H-bond	A	B	C
Hb1	1.81	2.11	2.16 ± 0.3
Hb2	1.81	1.71	1.74 ± 0.1
Hb3	1.91	2.45	2.40 ± 0.6
Hb4	2.55	1.73	1.76 ± 0.1
Hb5	1.66	1.72	1.75 ± 0.1
Hb6	1.74	1.74	1.18 ± 0.1
Hb7	1.94	1.91	1.95 ± 0.2
Hb8	1.89	1.89	1.93 ± 0.1
Hb9	1.78	1.82	1.87 ± 0.1
Hb10	1.69	1.80	1.77 ± 0.1
Hb11	1.82	2.02	2.07 ± 0.2
Hb12	2.16	2.26	2.36 ± 0.5

†Column A represents the H-bonds length in the minimized system (starting point of the MD simulation). Column B represents the H-bonds length in the MD averaged structure. Column C represents the average H-bonds length along the equilibrated MD simulation. Distances are in Å.

Multiple Sequence Alignments

Two sets of sequences were aligned. Set1 consists of the sequences of the proteins considered as the phylogenetic evolution of CDKs,[18] which are believed to be cdc2 *(Pneumocistis carinii)*, cdc2 *(Drosophila melanogaster)*, cdc2 *(Arabidopsis thaliana)*, cdk2 (*Xenopus laevis)*, cdk2 *(Carassius auratus)*, cdk2 *(Rattus norvegicus)*, and cdk2 *(Homo sapiens)* (see Table III of the Supplementary Material). Set2 is constituted by the sequences of the 11 CDKs (CDK1–CDK11) forming the CDK protein family.

Similarity alignments of the whole protein sequence of both sets were carried out using the utility Plotcon from the Emboss software suite[32] [see Fig. 9 (A1,B1)]. Multiple alignments of the AL sequence, delimited by DFG and APE motifs, were performed using Clustal W 1.81 program[33] [see Fig. 9(A2,B2)].

RESULTS AND DISCUSSION

MD Simulations

General feature of pCDK2, CDK2 and CDK2-II.

The overall folds are well conserved during the dynamics [Fig. 4(A–C)]. After ~3.5 ns, the RMSD of the backbone atoms of the entire systems fluctuates around a value of 1.6–1.8 Å (Table I). The RMSD of the (p)CDK2, Cyclin A,

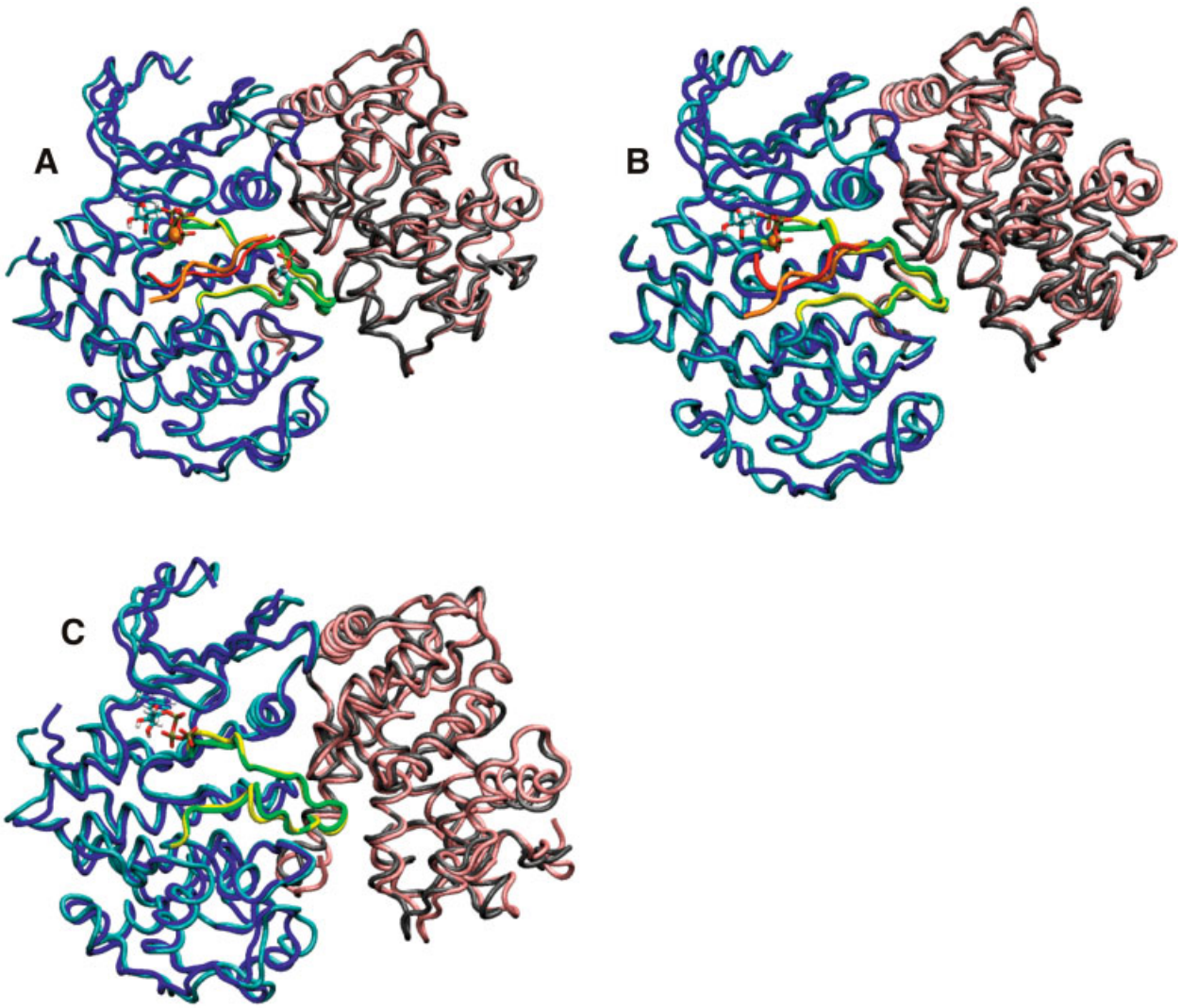


Fig. 4. MD averaged structures. Superposition of X-ray and the MD averaged structures of pCDK2 (**A**), CDK2 (**B**), and CDK2-II (**C**) structures. The X-ray structures of (p)CDK2, Cyclin A, AL, and substrate (A and B) are colored in blue, gray, green, and orange, respectively. Those of the MD structures are light-blue, pink, yellow, and red, respectively. The Mg^{2+} cation (A and B) is represented as a sphere; ATP and (p)Thr160 as sticks, colored based on the atom-type. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

and PSTAIRE helix fluctuates around 1.5–1.8, 1.1–1.3, and 0.4–0.5 Å, respectively (Table II).

pCDK2

The overall core of the protein is quite rigid; the RMSD of the active site region fluctuates only 0.4 (0.8) Å (values are given for backbone atoms, while the corresponding all atoms value is given in parenthesis), confirming the quality of catalytic site residues and Mg^{2+} coordination shell parameterization.[24] Also, the CDK2-Cyclin A interaction is well maintained, and conserved along the whole simulation, as shown by Figure 4(A). The RMSD of AL fluctuates around 0.7 (1.5) Å.

The calculated B-factors of pCDK2 compare well with experimental data[16] [Fig. 5(A)]. The only small discrepancy is the presence of a peak for L1 and L2 loops, in contrast to the experimental result, which shows two peaks.

The H-bond network formed by the pThr160 and the three coordinated guanidinium groups of each arginine (Arg50, Arg126, Arg150) residue is well maintained during the MD simulation. Specifically, pThr160 H-bonds to Hb1@Arg150, Hb2@Arg150, Hb5@Arg126, Hb6@Arg126, Hb9@Arg50, and Hb10@Arg50. The other six H-bonds are formed by each arginine side chain and backbone or side chain atoms of residues in the surrounding: Hb3@Arg150 and Hb4@Arg150 with backbone and side chain oxygen atoms of Glu268@CycA, respectively; Hb7@Arg126 and Hb8@Arg126 with backbone oxygen atom of Leu148@CDK2; Hb11@Arg50 and Hb12@Arg50 with backbone oxygen atom of Phe267@CycA. The H-bond network is shown in Figure 6(A, B), while the stability of each H-bond is reported in Table III.

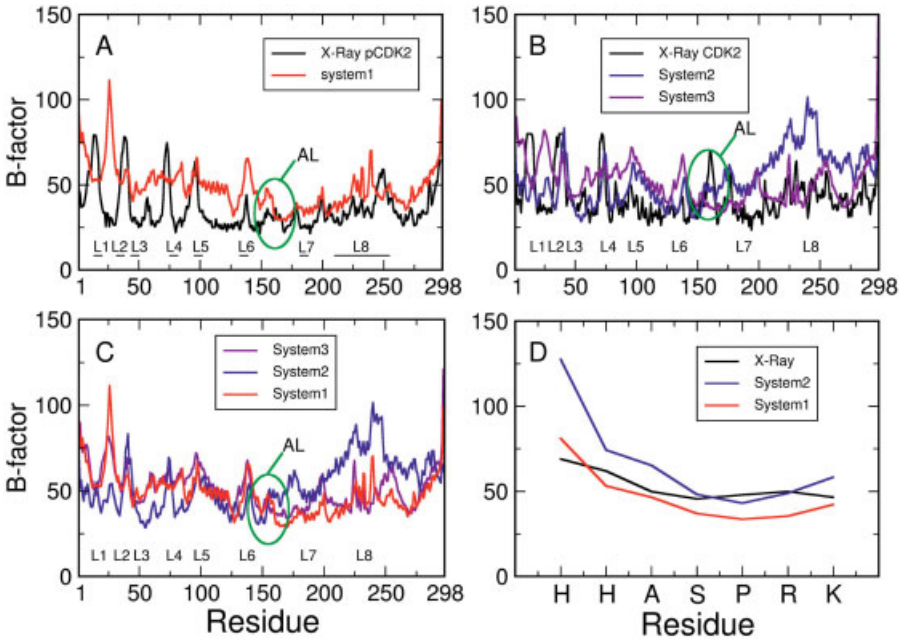


Fig. 5. Calculated and experimentally derived B-factors. **A:** Comparison of calculated B-factors per residue (Cα atoms) of pCDK2 (red) in pCDK2 and those experimentally derived by Brown et al.[16] (black). **B:** Calculated B-factors of CDK2 (blue) in CDK2 and in CDK2-II (violet) are compared with those experimentally derived (black) by Jeffrey et al.[11] **C:** Calculated B-factors comparison of the three systems. **D:** Calculated B-factors of Sub in pCDK2 (red) and CDK2 (blue) are compared with those experimentally derived by Brown et al.[16] (black). Residues are indicated in the single letter code, where S is the phosphorylable serine. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

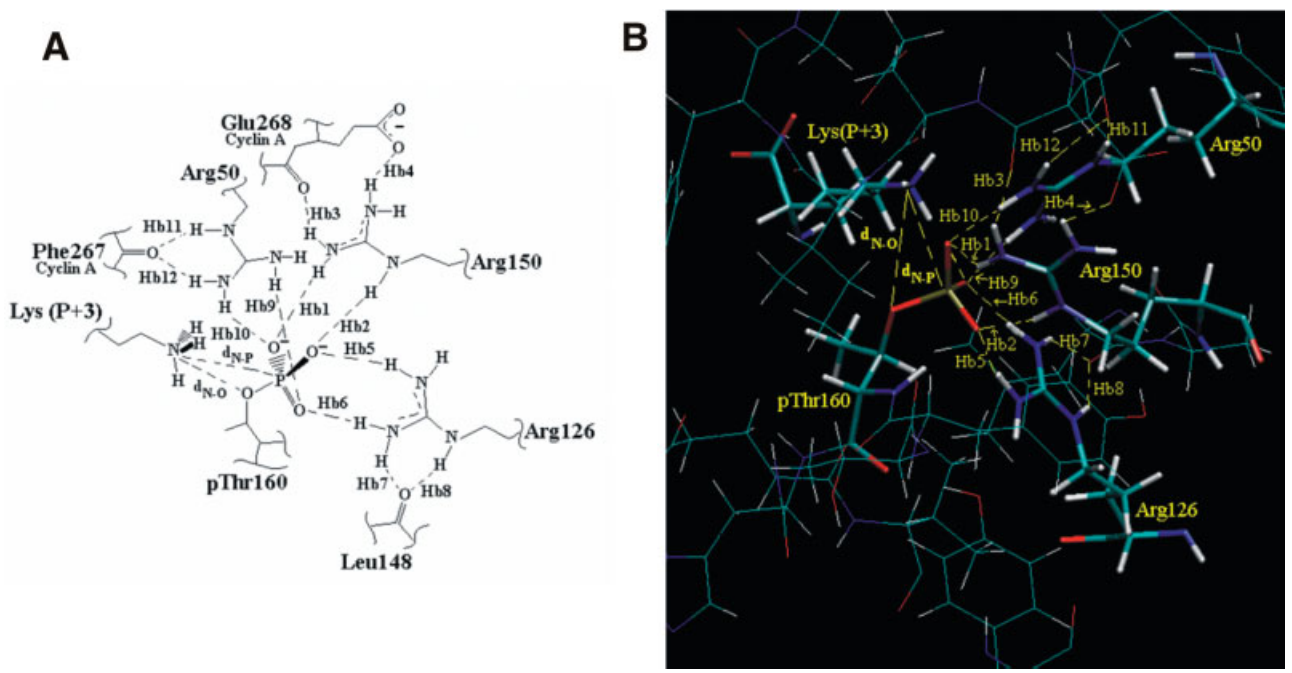


Fig. 6. H-bonds network around position 160 in pCDK2. **A:** The H-bonds interactions between the pThr160 phosphate group, P+3@Sub, and the three arginines (Arg50, Arg126, and Arg150) in pCDK2 are shown, along with the distances between the N_z terminal atom of P+3@Sub side chain and either the P atom of pThr160 ($d_{N\text{-}P}$) or the O_g atom of pThr160 side chain (d_{N-O}). **B:** H-bond network formed by pThr160, P+3@Sub, Arg50, Arg126, and Arg150 is represented in thick lines, the outer shell in thin lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

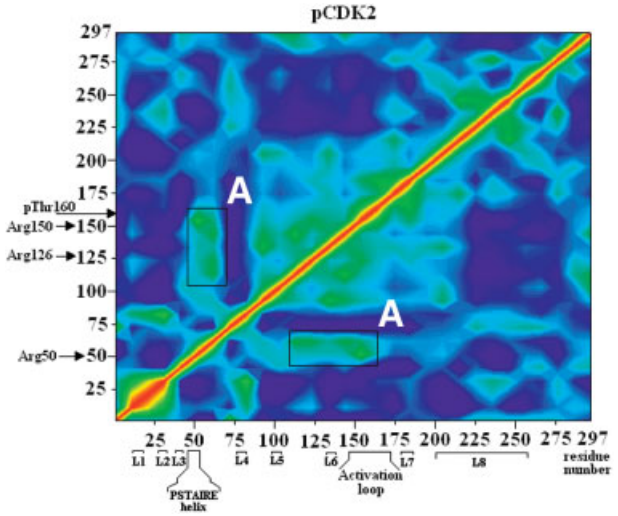


Fig. 7. Cross-correlation matrix of pCDK2. Residue numbers are indicated along the axes. Loop regions and residues involved in the pThr160 centered H-bond network are also reported. The matrix includes a region of interest (A). Colors range from red (1) to dark blue (−0.3). Green ranges from about 0.35 to 0.55. Labels referring to interesting regions of the protein are reported in both axes of the diagram. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The relative rigidity of the H-bonds network at the pThr160 site is also visible in the cross-correlated matrix, which exhibits an evident correlation of pThr160, Arg50, Arg126, and Arg150 (Fig. 7).

The model polypeptide substrate (Sub) shows fluctuations around 1.8 (2.6) Å. Importantly, the P+3@Sub-pThr160 interaction is very stable: both d_{N-P} and d_{N-O} distances (see Fig. 6) are maintained similar to their starting values (3.4 ± 0.1 and 3.5 ± 0.1 Å, respectively).

CDK2

The AL fluctuates around 1.1 (1.7) Å, showing a higher mobility than that observed in pCDK2. Consistently, whilst the calculated B-factors are overall rather similar to those of pCDK2 (Fig. 5), the B-factors of the AL have a different trend, underlining an increased flexibility for the unphosphorylated AL [Fig. 5(B,C)]. Also L8 presents an overall higher flexibility.

The larger flexibility of the loop is caused by the absence of the phosphate group on Thr160, which completely changes the H-bond network with the three guanidinium groups of each nearby arginine (Arg50, Arg126, and Arg150), in comparison with that present in pCDK2. Here, OH@Thr160 H-bonds only Arg126, while Arg150 and Arg50 interact with Glu268@CycA and Glu269@CycA, respectively. Moreover, the hole created by the lack of the phosphate is filled up by 5–6 water molecules, which thus partially hydrate the arginine residues no longer involved in electrostatic interactions with the lacking phosphate of Thr160 (Fig. 8).

Sub fluctuates significantly more than in pCDK2 (1.9 and 5.5 Å for the backbone and all atoms, respectively). Due to the absence of P+3@Sub-Thr160 interaction, P+3@Sub strongly changes its starting conformation, at the end pointing towards the bulk of the solvent. Consistently, the d_{N-O} distance fluctuates around as much as 13.4 ± 2.8 Å (Fig. 8).

CDK2-II

The ATP, which in this model is not coordinated to the Mg^{2+}, fluctuates around a considerably larger RMSD than in the other complexes (1.9 Å, Table II). The AL fluctuates around 1.2 (2.0) Å, again evidencing its minor stability in the opened and unphosphorylated form. The protein–protein interaction between CDK2 and Cyclin A is well maintained, and conserved along the whole simulation, as shown in Figure 4(C).

pThr160 is replaced by the carboxylate of Glu162 in the X-ray structure, which interacts mainly with Arg50 and Arg126, while Arg50 remains close by (~4 Å). In our MD simulation, the side chain is rather stable (RMSD value fluctuating around 0.8 Å). Also, Glu162 interacts either with the guanidinium group of Arg50 or that of Arg126 or that of Arg150, confirming the interactions present in the experimental structure. Also, 3–4 water molecules are included in the cavity partially hydrating the arginine residues not instantly involved in electrostatic interaction with Glu162. Importantly, in this model the Thr160 side chain points towards the solvent, showing a very stable orientation (all atoms RMSD value fluctuates around 0.8 Å). This H-bond network induces a favorable conformation of Thr160 for its subsequent phosphorylation by CAK.

Multiple Sequence Alignments

Our simulations show that the AL flexibility is correlated to an increase of the CDK2 enzymatic activity. In order to extend our results to all CDKs, MD calculations are here complemented by multiple sequence alignments. Thus, two sets of sequences were collected: set1 is formed by the sequences of the seven proteins constituting the phylogenetic evolution of CDKs; set2 is formed by the sequences of the eleven CDKs (CDK1-CDK11) of the CDKs protein family.

First, we have performed similarity sequence alignments of set1 [Fig. 9(A1)], then the same similarity sequence alignments of set2 [Fig. 9(B1)]. In both similarity plots, the AL sequence is one of the most conserved regions throughout the whole sequence, supporting its essential role in the enzymatic activation of CDKs. The negative peak throughout the AL sequence in the similarity plot in Figure 9(B1) is due to a not conserved insertion, also well visible in Figure 9(B2).

Multiple alignments of the AL sequences, delimited by DFG and APE motifs, further underline their very high conservation in both set1 and set2 [see Fig. 9(A2,B2)].

CONCLUSIONS

Fluctuations of protein residues and prosthetic groups play a crucial role in regulating protein functions.[34,35] It is now established that, although invaluable insight can be obtained by simplified static representations of proteins in terms of their average structures and B-factors, many

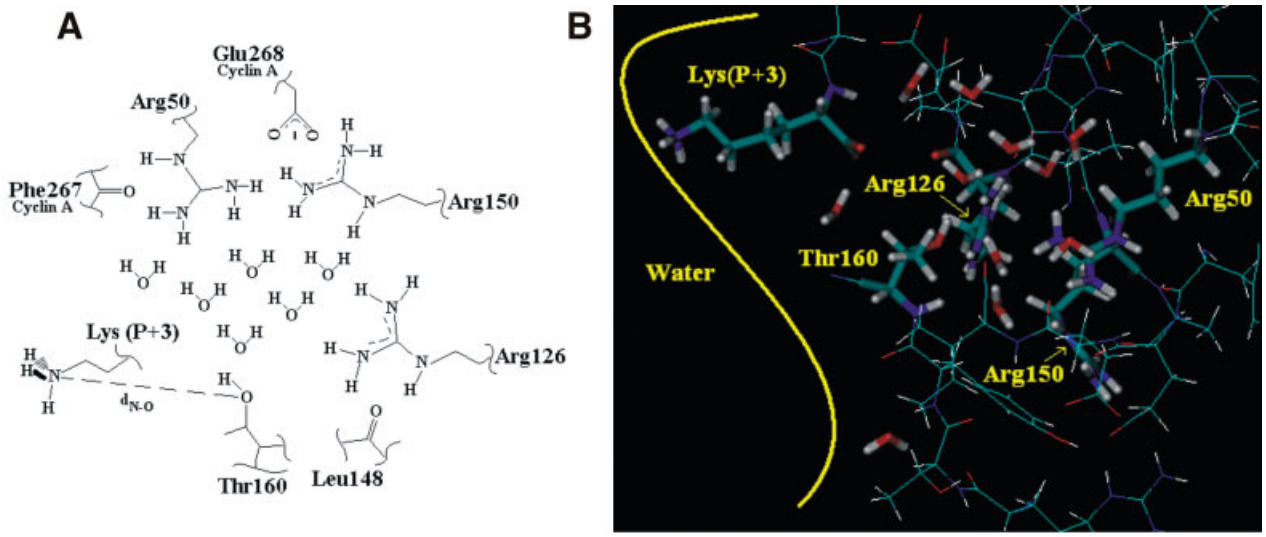
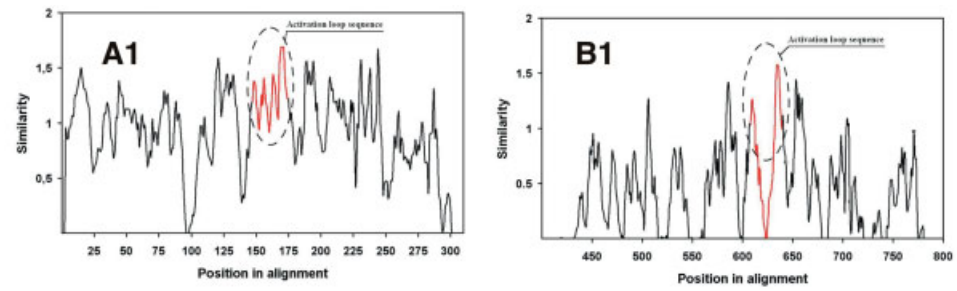
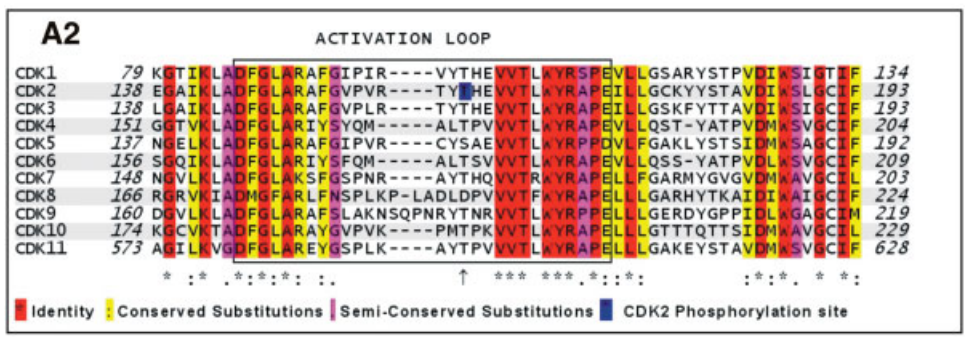


Fig. 8. **A:** The complex network of H-bonds, formed by the OH@Thr160 group, P+3@Sub, and the three arginines (Arg50, Arg126, and Arg150) in CDK2, is shown. Also the distances d_{N-O} (distance between the N_z terminal atom of P+3@Sub side chain and O_g atom of Thr160 side chain) is indicated. **B:** Thr160, P+3@Sub, Arg50, Arg126, and Arg150, and water molecules in the hole created by the removal of the phosphate group are represented in thick lines, the outer shell in thin lines. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]





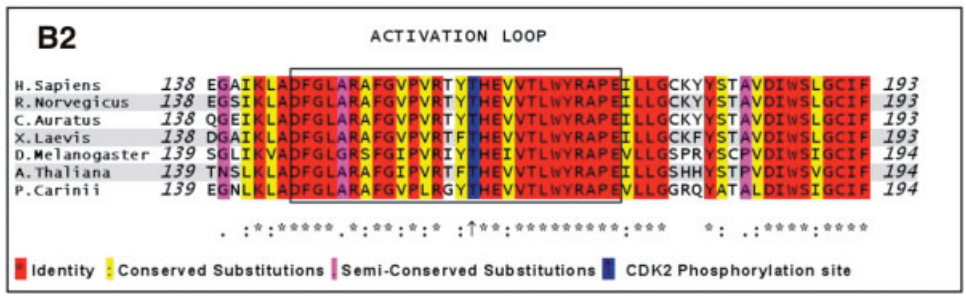


Fig. 9. Multiple alignments of CDKs. **A1:** Similarity plot of the whole sequence of the seven proteins considered as the phylogenetic evolution of CDKs. **B1:** Similarity plot of the whole sequence of the 11 CDKs constituting the CDKs protein family. In both A1 and B1, the AL region is highlighted and colored in red. **A2:** Multiple alignments of the only AL sequence of the proteins forming the phylogenetic evolution of CDKs. **B2:** Multiple alignments of the only AL sequence of the proteins forming the CDKs family. In both A2 and B2, the AL sequence is indicated by a black box. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

dangerous approximations are present in considering a unique fixed average structure.[34] In this respect, MD simulations can provide complementary and important information.

Here we have addressed this issue in the context of CDK2 performing a comparative MD study on three systems that differ mainly for the presence or absence of the phosphate group bound to Thr160, namely: systems pCDK2 and CDK2, based on a phosphorylated X-ray structure, and system CDK2-II based on an unphosphorylated X-Ray structure. Our goal has been to investigate the role of Thr160 for the flexibility of AL and CDK2 cellular partners, Cyclin A and peptide substrate.

Our calculations show that the overall folding of the models, including either pThr160 or Thr160, is well conserved. Nevertheless, we found a lower flexibility of the AL in the phosphorylated model (pCDK2), mostly because of the presence of charge interactions and H-bond network between pThr160 and Arg50, Arg126, and Arg150. These interactions are completely lost in the non-phosphorylated complex (CDK2) with the consequent increased mobility of the AL. The calculated B-factors of pCDK2 are in fairly good agreement with the experimental data, substantiating our findings. However, the AL is quite stable also in CDK2 and CDK2-II, thus explaining the basal enzymatic activity found experimentally for the partially activated CDK2/Cyclin A binary complex.[10] Hence, this result supports the role of Cyclin A in stabilizing the AL in the open conformation.

Our results also highlight that the pThr160 plays a role in substrate recognition and stability, through the interaction between the phosphate group and the basic P+3@Sub residue, as previously proposed by Brown et al.[16] In the CDK2-II simulation, Glu162 keeps its starting interactions with the arginine residues, while the Thr160 side chain points towards the solvent in favorable conformation for its subsequent phosphorylation by CAK.

In conclusion, our calculations offer insights on the molecular mechanism of the activation,[10] specifically on the role of dynamical fluctuations of crucial protein residues such as pThr160 and Arg50, Arg126 and Arg150 for the activity of the CDK2/Cyclin A/Substrate ternary complex.[10] Multiple sequence alignments, performed here, point to a very high conservation of the specific AL sequence in both series (Fig. 9), suggesting that the features found here are rather general for the CDK family.

These results, along with structural information,[36] might be useful in designing specific CDKs inhibitors exploiting high-affinity interactions based on protein flexibility.[37]

ACKNOWLEDGMENTS

We thank INFM and MIUR-COFIN for financial support, and the supercomputer center CINECA (Bologna, Italy) for computing time. Also, we thank Cristian Micheletti for providing us the coarse-grain model recently described in Micheletti et al.[31]

REFERENCES

1. Morgan DO. Principles of CDK regulation. Nature 1995;374:131–134.
2. Harper JW, Adams PD. Cyclin-dependent kinases. Chem Rev 2001;101:2511–2526.
3. Morgan DO. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol 1997;13:261–291.
4. Knockaert M, Greengard P, Meijer L. Pharmacological inhibitors of cyclin-dependent kinases. Trends Pharmacol Sci 2002;23:417–425.
5. Fisher RP, Morgan DO. A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. Cell 1994;78: 713–724.
6. Espinoza FH, Farrell A, Erdjument-Bromage H, Tempst P, Morgan DO. A cyclin-dependent kinase-activating kinase (CAK) in budding yeast unrelated to vertebrate CAK. Science 1996;273: 1714–1717.
7. Kaldis P, Sutton A, Solomon MJ. The Cdk-activating kinase (CAK) from budding yeast. Cell 1996;86:553–564.
8. Pavletich NP. Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. J Mol Biol 1999;287:821–828.
9. De Bondt HL, Rosenblatt J, Jancarik J, Jones HD, Morgan DO, Kim SH. Crystal structure of cyclin-dependent kinase 2. Nature 1993;363:595–602.9.
10. Russo AA, Jeffrey PD, Pavletich NP. Structural basis of cyclin-dependent kinase activation by phosphorylation. Nat Struct Biol 1996;3:696–700.
11. Jeffrey PD, Russo AA, Polyak K, Gibbs E, Hurwitz J, Massague J, Pavletich NP. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature 1995;376:313–320.
12. Fernandez-Fuentes N, Hermoso A, Espadaler J, Querol E., Aviles FX, Oliva B. Classification of common functional loops of kinase super-families. Proteins 2004;56(3):539–555.
13. Brown NR, Noble ME, Lawrie AM, Morris MC, Tunnah P, Divita G, Johnson LN, Endicott JA. Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity. J Biol Chem 1999;274:8746–8756.
14. Johnson LN, Lewis RJ. Structural basis for control by phosphorylation. Chem Rev 2001;101:2209–2242.
15. Park H, Yeom MS, Lee S. Loop flexibility and solvent dynamics for the selective inhibition of cyclin-dependent kinase 4: comparative molecular dynamics simulation studies of CDK2 and CDK4. Chem Bio Chem 2004;5;1662–1672.
16. Brown NR, Noble ME, Endicott JA, Johnson LN. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. Nat Cell Biol 1999;1:438–443.
17. Stevenson-Lindertv LM, Fowler P, Lew J. Substrate Specificity of CDK2-Cyclin A. J Biol Chem 2003;278;50956–50960.
18. Krylov DM, Nasmyth K, Koonin EV. Evolution of eukaryotic cell cycle regulation: stepwise addition of regulatory kinases and late advent of the CDKs. Curr Biol 2003;13:173–177.
19. Cavalli A, De Vivo M, Recanatini M. Density functional study of the enzymatic reaction catalyzed by a cyclin-dependent kinase. Chem Commun (Camb) 2003:1308–1309.
20. Cohen P. Protein kinases: the major drug targets of the twenty-first century? Nat Rev Drug Discov 2002;1:309–315.
21. Dancey J, Sausville EA. Issues and progress with protein kinase inhibitors for cancer treatment. Nat Rev Drug Discov 2003;2:296–313.
22. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA. AMBER 7. In. San Francisco: University of California; 2002.
23. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
24. Cavalli A, Dezi C, Folkers G, Scapozza L, Recanatini M. Three-dimensional model of the cyclin-dependent kinase 1 (CDK1): Ab initio active site parameters for molecular dynamics studies of CDKS. Proteins 2001;45:478–485.
25. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model. J Phys Chem 1993;97:10269–10280.
26. Banci L, Schroder S, Kollman PA. Molecular dynamics character-

ization of the active cavity of carboxypeptidase A and some of its inhibitor adducts. Proteins 1992;13:288–305.
27. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–341.
28. Darden T, York D, Pedersen L. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. J Chem Phys 1993;98: 10089–10094.
29. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103: 8577–8593.
30. Berendsen HJC, Postma JPM, Van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
31. Micheletti C, Carloni P, Maritan A. Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models. Proteins 2004;55:635–645.
32. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 2000;16:276–277.
33. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
34. Allen TW, Anderson OS, Roux B. On the importance of atomic fluctuations, protein flexibility, and solvent in ion permeation. J Gen Physiol 2004;124;679–690.
35. Piana S, Carloni P, Parrinello M. Role of confirmational fluctuations in the enzymatic reaction of HIV-1 protease. J Mole Bio 2002;319(2):567–583.
36. Noble ME, Endicott JA, Johnson LN. Protein kinase inhibitors: insights into drug design from structure. Science 2004;303:1800–1805.
37. Teague SJ. Implications of protein flexibility for drug discovery. Nat Rev Drug Discov 2003;2:527–541.