J|A|C|S ARTICLES

Published on Web 12/09/2008

Protein Conformational Transitions: The Closure Mechanism of a Kinase Explored by Atomistic Simulations

Anna Berteotti,[†] Andrea Cavalli,[*,‡,§] Davide Branduardi,[||] Francesco Luigi Gervasio,[*,||] Maurizio Recanatini,[‡] and Michele Parrinello[||]

Scuola Normale Superiore, Piazza dei Cavalieri, I-56126 Pisa, Italy, Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy, Unit of Drug Discovery and Development, Italian Institute of Technology, Via Morego 30, I-16163 Genoa, Italy, and Computational Sciences, Department of Chemistry and Applied Biosciences, ETH Zurich, USI Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

Received August 29, 2008; E-mail: andrea.cavalli@unibo.it; francesco.gervasio@phys.chem.ethz.ch

Ⓦ **This paper contains enhanced objects available on the Internet at http://pubs.acs.org/jacs.**

Abstract: Kinase large-scale conformational rearrangement is an issue of enormous biological and pharmacological relevance. Atomistic simulations able to capture the dynamics and the energetics of kinase large-scale motions are still in their infancy. Here, we present a computational study in which the atomistic dynamics of the "open-to-closed" movement of the cyclin-dependent kinase 5 (CDK5) have been simulated. Simulations were carried out using a new sampling method that is able to find the lowest free-energy channel between an initial state and a final state. This large-scale movement has a two-step mechanism: first, the αC-helix rotates by ∼45°, allowing the interaction between Glu51 and Arg149; then the CDK5 activation loop refolds to assume the closed conformation. We have also estimated the free-energy profile associated with the global motion and identified a CDK5 intermediate, which could be exploited for drug-design purposes. Our new sampling method turned out to be well-suited for investigating at an atomistic level the energetics and dynamics of kinase large-scale conformational motions.

Introduction

Kinases represent a vast protein family involved in vital cellular pathways and responsible for several biochemical functions.[1] Understanding the conformational dynamics of kinases[2] may represent a bridge between their structure and function.[3] Indeed, the activation of many cellular mechanisms requires kinases to undergo large-scale intrinsic motions, as these proteins can adopt at least two extreme conformations, an "open" state that is maximally active and a "closed" state that shows minimal activity.[2] Structural insights into the open and closed states of kinases have been gained from a number of case studies, where two or more crystal structures of the same protein in different conformations were determined by X-ray experiments.[3,4] However, an understanding of the energetic profile and the atomistic dynamics of the transition between the open and closed states is still in its infancy. Moreover, knowledge of the structures of possible intermediate conformations sampled during the transition is still very limited. Therefore, X-ray data must be complemented with other approaches able to capture the dynamics of large-scale conformational movements.[5] Recently, conformational transitions have been measured by NMR relaxation experiments with substrate analogues[6−9] or during catalysis[10,11] as well as by single-molecule fluorescence resonance energy transfer (FRET) measurements.[12−17]

[†] Scuola Normale Superiore.
[‡] University of Bologna.
[§] Italian Institute of Technology.
[||] ETH Zurich.

(1) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912–1934.
(2) Huse, M.; Kuriyan, J. *Cell* **2002**, *109*, 275–282.
(3) Levinson, N. M.; Kuchment, O.; Shen, K.; Young, M. A.; Koldobskiy, M.; Karplus, M.; Cole, P. A.; Kuriyan, J. *PLoS Biol.* **2006**, *4*, e144.
(4) Morgan, D. O.; De Bondt, H. L. *Curr. Opin. Cell Biol.* **1994**, *6*, 239–246.
(5) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D. *Nature* **2007**, *450*, 838–844.
(6) Williams, J. C.; McDermott, A. E. *Biochemistry* **1995**, *34*, 8309–8319.
(7) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313*, 1638–1642.
(8) Palmer, A. G. *Chem. Rev.* **2004**, *104*, 3623–3640.
(9) Cui, Q.; Karplus, M. *Adv. Protein Chem.* **2003**, *66*, 315–372.
(10) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. *Nature* **2005**, *438*, 117–121.
(11) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. *Nat. Struct. Mol. Biol.* **2004**, *11*, 945–949.
(12) Blanchard, S. C.; Gonzalez, R. L.; Kim, H. D.; Chu, S.; Puglisi, J. D. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1008–1014.
(13) Ha, T.; Ting, A. Y.; Liang, J.; Caldwell, W. B.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 893–889.
(14) Myong, S.; Stevens, B. C.; Ha, T. *Structure* **2006**, *14*, 633–643.
(15) Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wohrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1655–1660.
(16) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743–747.
(17) Zhang, Z.; Rajagopalan, P. T.; Selzer, T.; Benkovic, S. J.; Hammes, G. G. *Proc. Natl. Acad. Sci. U.S A.* **2004**, *101*, 2764–2769.

 10.1021/ja806846q CCC: $40.75

In this scenario, atomistic simulations could greatly contribute to linking flexibility and function.[5,18−20] Although conventional molecular dynamics (MD) studies in explicit water solution can access only a limited time scale (∼100 ns), while large-scale conformational rearrangements occur on the microsecond-to-millisecond time scale, several schemes have been developed to overcome this drawback, most of them based on a coarse-grained representation of the system.[21−24] This approach can be very useful; however, obtaining a good coarse-grained representation can be difficult, and transferability is a major concern. Recently, we have proposed a scheme[25] that is able to reconstruct the lowest free-energy path that connects given initial and final states without having to renounce a fully atomistic description. This method has been successfully applied to study the conformational movements of small peptides,[25] but studies on large-scale conformational movements of protein systems in explicit solvent have not been reported to date.

Here we investigate kinase plasticity by studying the energetics and dynamics of the intrinsic conformational motions experienced by cyclin-dependent kinase 5 (CDK5) during the "open-to-closed" transition. Cyclin-dependent kinases (CDKs) are key players in the machinery that drives the progression and timing of the eukaryotic cell cycle,[26] and malfunctioning of CDKs is the basis of several diseases.[27−30] To be fully active, CDKs require binding of homologous protein activators known as cyclins and phosphorylation of a conserved threonine of the activation loop. Similarly to other kinases, during the open-to-closed transition, CDKs undergo a large-scale intrinsic rearrangement that results in rotation of an α-helix by almost 90° and relocation of the activation loop in a movement of ∼25 Å.[31,32]

CDK5[33] is a member of the CDK protein family, and it does not seem to be involved in cell-cycle regulation. Moreover, it is selectively activated by p35 or p39, whose expression is limited to neurons and a few other cell types.[34] As a consequence, CDK5 is implicated in neuronal development and maintenance of adult neuronal architecture,[35] and its deregulation has been associated with a number of neurodegenerative diseases, such as Alzheimer's disease,[36] Parkinson's disease,[37] amyotrophic lateral sclerosis,[38] Niemann−Pick type C disease,[39] and ischemia.[40] Furthermore, CDK5 does not need the phosphorylation of the activation loop to be fully active. Actually, in the X-ray structure of the CDK5/p25 complex,[41] the structural requirement for phosphorylation is bypassed by an extended network of interactions between p25 and the CDK5 activation loop. The absence of the phosphorylation step renders CDK5 particularly well-suited for investigating kinase large-scale intrinsic motions, and moreover, it represents a very interesting target for drug discovery.

Methods

Homology Modeling. The structures of open CDK5 and the CDK5/p25 complex were obtained from the crystal structure of CDK5 refined at 2.2 Å resolution in a complex with roscovitine (PDB entry 1UNL)[42] by removing the inhibitor and the activator p25 and replacing Asn144 with an Asp, in order to have an active protein. The search for a reactive path starts from the definition of an initial (open) state and a final (closed) state. Since the crystal structure of CDK5 in the closed state is not available, we built a chimeric structure in which the αC-helix and T-loop were obtained by homology modeling of CDK5 on the closed structure of CDK2 (PDB entry 1FIN[43]). These regions show sequence identities (74 and 85% for the αC-helix and the T-loop, respectively) that are even higher than that related to the entire protein (60%). The remaining residues were taken from the original structure of CDK5 in the open conformation (PDB entry 1UNL[42]). The homology model was generated with the program Modeler version 7.0.[44] The structures thus obtained were ranked on the basis of the internal scoring function of Modeler and validated with PROCHECK.[45] According to the PROCHECK analysis, the homology-built CDK5 closed conformation turned out to be a very high-quality model, with an overall G-factor of −0.04. The Ramachandran plot[46] related to this model showed that more than 90% of the residues were located in the most favored regions [see Supporting Figure (SF) 6 in the Supporting Information].

Molecular Dynamics. Preliminary MD simulations of the open state, the closed state described above, and the CDK5/p25 complex were carried out using the GROMACS 3.2.1 suite[47] and the OPLS-AA force field.[48] Water solvation was described by immersion of the open and closed structures in ∼13 000 TIP3P waters, while for the complex, a larger number (∼ 23 000) was necessary. Neutrality was ensured by adding one Cl^- to the two isolated proteins and two Na^+ ions to the complex. Periodic boundary conditions and the particle mesh Ewald method (to account for long-range electrostatic interactions) were used throughout. Bonds involving hydrogens were constrained using the SHAKE algorithm,[49] and

(18) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585–590.
(19) Norberg, J.; Nilsson, L. *Q. Rev. Biophys.* **2003**, *36*, 257–306.
(20) Barrett, C. P.; Noble, M. E. *J. Biol. Chem.* **2005**, *280*, 13993–14005.
(21) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.
(22) Yang, S.; Roux, B. *PLoS Comput. Biol.* **2008**, *4*, e1000047.
(23) Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. *J. Mol. Biol.* **2007**, *366*, 1661–1671.
(24) Lu, Q.; Wang, J. *J. Am. Chem. Soc.* **2008**, *130*, 4772–4783.
(25) Branduardi, D.; Gervasio, F. L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.
(26) Morgan, D. O. *Annu. Rev. Cell Dev. Biol.* **1997**, *13*, 261–291.
(27) Lee, Y. M.; Sicinski, P. *Cell Cycle* **2006**, *5*, 2110–2114.
(28) Geyer, J. A.; Prigge, S. T.; Waters, N. C. *Biochim. Biophys. Acta* **2005**, *1754*, 160–170.
(29) Schang, L. M. *Curr. Drug Targets: Infect. Disord.* **2005**, *5*, 29–37.
(30) Monaco, E. A., 3rd; Vallano, M. L. *Front Biosci.* **2005**, *10*, 143–159.
(31) Pavletich, N. P. *J. Mol. Biol.* **1999**, *287*, 821–828.
(32) Obaya, A. J.; Sedivy, J. M. *Cell. Mol. Life Sci.* **2002**, *59*, 126–142.
(33) Ishiguro, K.; Kobayashi, S.; Omori, A.; Takamatsu, M.; Yonekura, S.; Anzai, K.; Imahori, K.; Uchida, T. *FEBS Lett.* **1994**, *342*, 203–208.
(34) Dhavan, R.; Tsai, L. H. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 749–759.
(35) Cheung, Z. H.; Fu, A. K.; Ip, N. Y. *Neuron* **2006**, *50*, 13–18.
(36) Cruz, J. C.; Tsai, L. H. *Trends Mol. Med.* **2004**, *10*, 452–458.
(37) Smith, P. D.; Crocker, S. J.; Jackson-Lewis, V.; Jordan-Sciutto, K. L.; Hayley, S.; Mount, M. P.; O'Hare, M. J.; Callaghan, S.; Slack, R. S.; Przedborski, S.; Anisman, H.; Park, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13650–13655.
(38) Nguyen, M. D.; Julien, J. P. *Neurosignals* **2003**, *12*, 215–220.
(39) Hallows, J. L.; Iosif, R. E.; Biasell, R. D.; Vincent, I. *J. Neurosci.* **2006**, *26*, 2738–2744.
(40) Camins, A.; Verdaguer, E.; Folch, J.; Pallas, M. *CNS Drug Rev.* **2006**, *12*, 135–148.
(41) Tarricone, C.; Dhavan, R.; Peng, J.; Areces, L. B.; Tsai, L. H.; Musacchio, A. *Mol. Cell* **2001**, *8*, 657–669.
(42) Mapelli, M.; Massimiliano, L.; Crovace, C.; Seeliger, M. A.; Tsai, L. H.; Meijer, L.; Musacchio, A. *J. Med. Chem.* **2005**, *48*, 671–679.
(43) Jeffrey, P. D.; Russo, A. A.; Polyak, K.; Gibbs, E.; Hurwitz, J.; Massague, J.; Pavletich, N. P. *Nature* **1995**, *376*, 313–320.
(44) Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.
(45) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(46) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95–99.
(47) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
(48) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
(49) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–342.

the time step for all of the simulations was 2 fs. A steepest-descent minimization and thermalization scheme was applied to all of the initial structures. The systems were heated from 0 to 100 K in 50 ps and to 300 K in 200 ps, keeping the Cα atoms fixed in their original positions. Next, all of the constraints were lifted, and the equilibration was continued in the isobaric−isothermal ensemble with Nose−Hoover thermostats[50] and the Parrinello−Rahman barostat[51] for ∼1 ns. For the open state and CDK5/p25 complex, after the thermalization scheme we performed ∼10 ns of MD on the canonical ensemble. In these simulations, the root-mean-square deviation (rmsd) of the protein and Cα showed good thermal stability, with a value of 1.5−2 Å. These runs were then analyzed by principal component analysis (see SF 2 in the Supporting Information).

Building the Guess Path. After MD thermalization, the open and closed conformations of CDK5 were submitted to the Molmov morphing server.[52] Starting from our initial and final structures, we obtained five interpolated intermediates. Since the bond lengths are kept fixed to their experimental values and an energy minimization is employed, the resulting sequence of structures is only a rough approximation to the reaction coordinate. In SF 7 in the Supporting Information, the different conformations of the αC-helix and the T-loop of the five CDK5 intermediate conformers (as output from Molmov) are shown along with the open and closed states of the enzyme. Each of the five morphing conformations was then hydrated and thermalized at 300 K (see above), and the last snapshots were used for generating the initial guess path. Notably, some morphing conformations were quite stable during the thermalization, while others showed a rapid increase of the rmsd of the Cα atoms. The seven conformations (the open state, the closed state, and the five morphing conformers) were used to generate the guess for subsequent path optimizations and metadynamics-based computations, which were carried out using the path collective variables (PCVs) s and z described below.

Metadynamics with Path Collective Variables. Following our previous work,[25] we defined two variables, $s(R)$ and $z(R)$, that are able to describe the position of a point in configurational space R relative to a preassigned path. In the present case, the preassigned path coincides with the morphing path. The variables s and z can be written as:

$$s(R) = \frac{1}{P-1} \frac{\sum_{i=1}^{P} (i-1) e^{-\lambda[R-R(i)]^2}}{\sum_{i=1}^{P} e^{-\lambda[R-R(i)]^2}} \quad (1)$$

$$z(R) = -\frac{1}{\lambda} \ln\left(\sum_{i=1}^{P} e^{-\lambda[R-R(i)]^2} \right) \quad (2)$$

where i is a discrete index ranging from 1 to P, where P is the number of conformations in the "template" path. In this way, the variable $s(R)$ provides the progress of the dynamics along the template path, while $z(R)$ provides the distance from the template path itself. In the present case, $[R - R(i)]^2$ is calculated as the mean square displacement after optimal alignment.[53] The choice of the coordinates of atoms to be included in the description is far from trivial: a wrong choice can turn into a loss of performance. An in-depth analysis of several trial trajectories in which we used only the Cα and several N atoms that belong to the αC-helix (30 to 60) and T-loop (145 to 165) showed that the side chains of Lys33, Arg50, Glu51, Asp144, and Arg149 also must be taken into account to effectively simulate the overall movement. For the discrete representation to be meaningful, we chose the nodal points $R(i)$ to be as equidistant as possible and the value of λ to be comparable to the inverse of the mean square displacement between successive frames.

We then ran a metadynamics[54] simulation to reconstruct the free-energy surface (FES) as a function of s and z. We used metadynamics in the direct version,[55] as implemented in NAMD.[56] The height of the Gaussians was set to 0.3 kcal/mol and the addition frequency to 1 ps. The width of the Gaussians was not fixed but was determined by the fluctuations in s and z measured over a time interval of 2 ps. In order to avoid integration problems with too-narrow Gaussians, a lower-bound limit of 0.03 in the appropriate unit was set for each coordinate.

We first optimized the path with $P = 7$ using the procedure described in ref 57. It was clear, however, that this discretization was not sufficient to describe the transition with the necessary resolution. We thus increased P to 18 by selecting additional intermediate configurations from among the small-z configurations obtained during the metadynamics. We chose these configurations such that they were equally spaced with a distance of 1.31 Å. The 18 frames were used to define a reference path using $\lambda = 1.33$ Å^{-2} in eqs 1 and 2. The path was not further improved, but it was sufficiently accurate for the study of FES(s, z) to provide us with reliable information on the open-to-closed transition and its associated free-energy profile.

Results and Discussion

The structural organization of CDK5 is reported in Figure 1 and hereafter briefly summarized. CDK5 exhibits the classical bilobal kinase fold, where the N-terminal domain (N-lobe) is composed mainly of the β-sheet and one α-helix (the αC-helix) (Figure 1a). The C-terminal domain (C-lobe) is predominantly α-helical and is linked to the N-terminal by a flexible hinge (81 to 84 aa). Between the lobes there is a deep cleft that natively binds ATP. The ceiling of the ATP-binding site, named the G-loop (11 to 18 aa), is quite flexible, being rich in glycines.[58] Two structural elements are involved in the interaction with the activator[41] (p25) and in the catalytic function: the αC-helix, whose sequence is well-conserved in the CDK family, and the activation loop (T-loop), which binds the two terminal lobes (145 to 165 aa). The latter belongs to the activation segment, which is shown in detail in Figure 1b. This is a structurally well-characterized portion of kinases that begins with the conserved DFG tripeptide (144−146 following the CDK5 numbering) and comprises the magnesium-binding loop (143−148), β9 (149−150), the activation loop (T-loop in CDKs) (151−160), and the P+1 loop (161−169), ending with a second conserved tripeptide motif [APE in general but PPD in CDK5 (170−172)]. In the C-lobe there is also another fundamental loop, called the catalytic loop (C-loop, Figure 1a), that is involved in the reaction catalysis (122−131 aa).

Closure Movement of CDK5: A Two-Step Mechanism. In a preliminary calculation, we assessed the effect of p25 on the stability of the open form of CDK5. In SF 1 in the Supporting Information, the rmsd during a 10 ns MD simulation is reported for the CDK5/p25 complex (SF 1A) and for CDK5 alone (SF 1B). It can be seen that CDK5 in the complex with its activator p25 (SF 1A) is more rigid than CDK5 alone (SF 1B) and that

(50) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
(51) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
(52) Krebs, W. G.; Gerstein, M. *Nucleic Acids Res.* **2000**, *28*, 1665–1675.
(53) Kearsley, S. K. *Acta Crystallogr., Sect. A* **1989**, *45*, 628–635.
(54) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.
(55) Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. *J. Phys. Chem. B* **2005**, *109*, 6714–6721.
(56) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
(57) Henkelman, G. *J. Chem. Phys.* **2000**, *113*, 9901–9904.
(58) Morgan, D. O. *Nature* **1995**, *374*, 131–134.

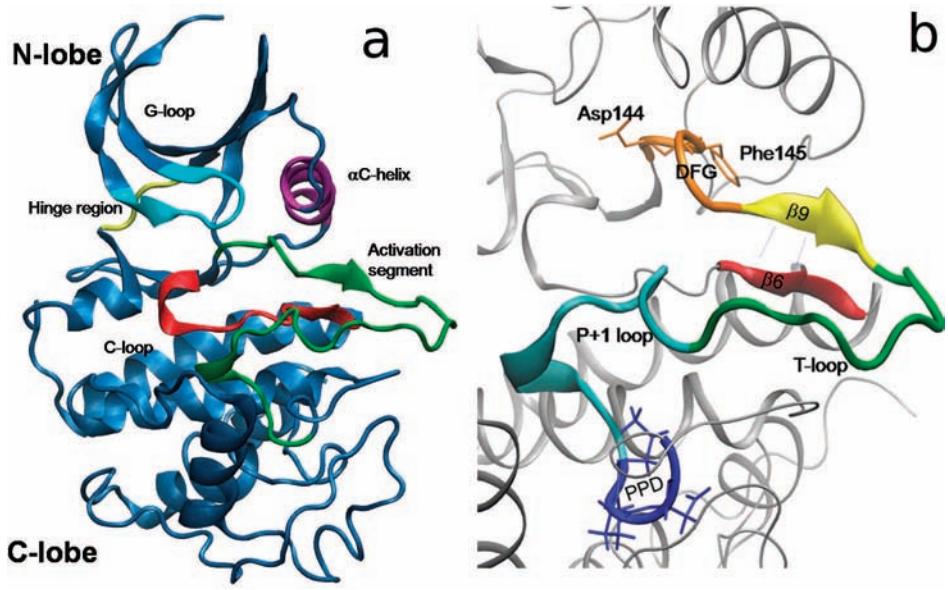

Figure 1. Three-dimensional structure of CDK5 in the open state. (a) Classical bilobal structure of kinase: (cyan) the G-loop (aa 11−18); (violet) the αC-helix (44−57); (yellow) the hinge region between the the N- and C-lobes of the protein (81−84); (red) the C-loop (122−131); (green) the activation segment (144−171). (b) Details of the activation segment: (orange) the DFG-in conformation, in which the Asp144 is pointing inside the cleft; the interaction between (yellow) β9 and (red) β6; (green) the T-loop; (cyan) the P+1 loop; (blue) PPD, the C-terminal part of the activation segment.

most of the flexibility of uncomplexed CDK5 comes from the αC-helix and the T-loop (blue trace in SF 1B). Similar results were obtained by principal component analysis (SF 2 in the Supporting Information). The stabilizing effect of the activator suggests that its role is to favor the open structure. This conclusion is strengthened by the observed dependence on activator concentration.[59] We simulated the closure process in the absence of the activator, which is a necessary step toward a full understanding of the enzyme functioning.

To determine the closing path (see Methods), we used metadynamics together with the method of Branduardi et al.[25] Although the initial guess path was constructed using a standard bioinformatics tool,[52] the final optimized path was very different from the initial guessed one.

Our simulations show that the process takes place in two steps: in the first step, the αC-helix rotates by ∼45° (Figure 2A,B), and in the second step, the T-loop closes the catalytic cleft and the αC-helix completes its rotation (Figure 2C,D). The first event leading to the rotation of the αC-helix is the breaking of the salt bridge between Lys33 and Glu51, which plays a key role in the kinase activation mechanism[60] (Figure 2A,B). The role of this salt bridge is to anchor the αC-helix in its open position. Solvation and the formation of a new salt bridge with Arg149, which is part of the activation segment, enable a partial rotation of the αC-helix. In the open conformation, Lys33 and Glu51 interact strongly and are assisted in stabilizing the αC-helix at its position by the close proximity of Asp144. The breaking of this bond takes place in a concerted fashion. Arg149 approaches Glu51, thus weakening the Lys33−Glu51 bond; this leads to a bond switch between Lys33−Glu51 and Lys33−Asp144. The end result is that the Lys33−Glu51 bond is disrupted and a new salt bridge, Glu51−Arg149, is formed, allowing the partial rotation of the αC-helix. It is also important to underline the role of the water in this first transition. In SF 3 in the Supporting Information, the reported changes in water coordination show that the breaking of the bond is assisted by the solvation of Glu51.

The DFG motif did not experience any in−out conformational movement, as observed in the crystal structures of CDK2 and CDK7 but at variance with CDK6, which can adopt a DFG-out conformation when it is bound to the inhibitor Ink4.[61]

In the second step, the αC-helix completes its rotation while the T-loop refolds to assume the final closed conformation. During the entire process, Arg149 and Glu51 remain bound (Figure 2C,D). Also in this second stage, Arg149 plays a major role by helping the T-loop to rearrange and maintain a connection between the C- and N-lobes. The mechanism that we have found by optimizing the path is closely related to one of the possible mechanisms described in ref 22 for the homologous Src kinase. In that paper, the authors used coarse-grained molecular representations and found that the closed-to-open transition can occur in two possible ways. The first pathway is similar to what we found in CDK5 and implies the opening of the activation loop followed by rotation of the αC-helix. Also, the switching of an electrostatic network, in particular between the highly conserved residues K295−E310 and E310−R409, plays a fundamental role in their case, as previously noted in ref 62 and recently confirmed in ref 63. On the contrary, we did not find evidence for the partial unfolding mechanism described in ref 22.

An added bonus of our atomically detailed pathway is that it shows the importance of hydration/dehydration of key residues during the restructuring of T-loop. Initially, the loop is relatively dehydrated. Halfway through the transition, the T-loop hydration reaches its maximum, being exposed to the solvent. At the end, water molecules must be expelled from the cavity, and T-loop hydration is reduced (SF 4 in the Supporting Information). After

(59) Hisanaga, S.; Saito, T. *Neurosignals* **2003**, *12*, 221–229.
(60) Nolen, B.; Taylor, S.; Ghosh, G. *Mol. Cell* **2004**, *15*, 661–675.
(61) Russo, A. A.; Tong, L.; Lee, J. O.; Jeffrey, P. D.; Pavletich, N. P. *Nature* **1998**, *395*, 237–243.
(62) Ozkirimli, E.; Post, C. B. *Protein Sci.* **2006**, *15*, 1051–1062.
(63) Ozkirimli, E.; Yadav, S. S.; Miller, W. T.; Post, C. B. *Protein Sci.* **2008**, *17*, 1871–1880.

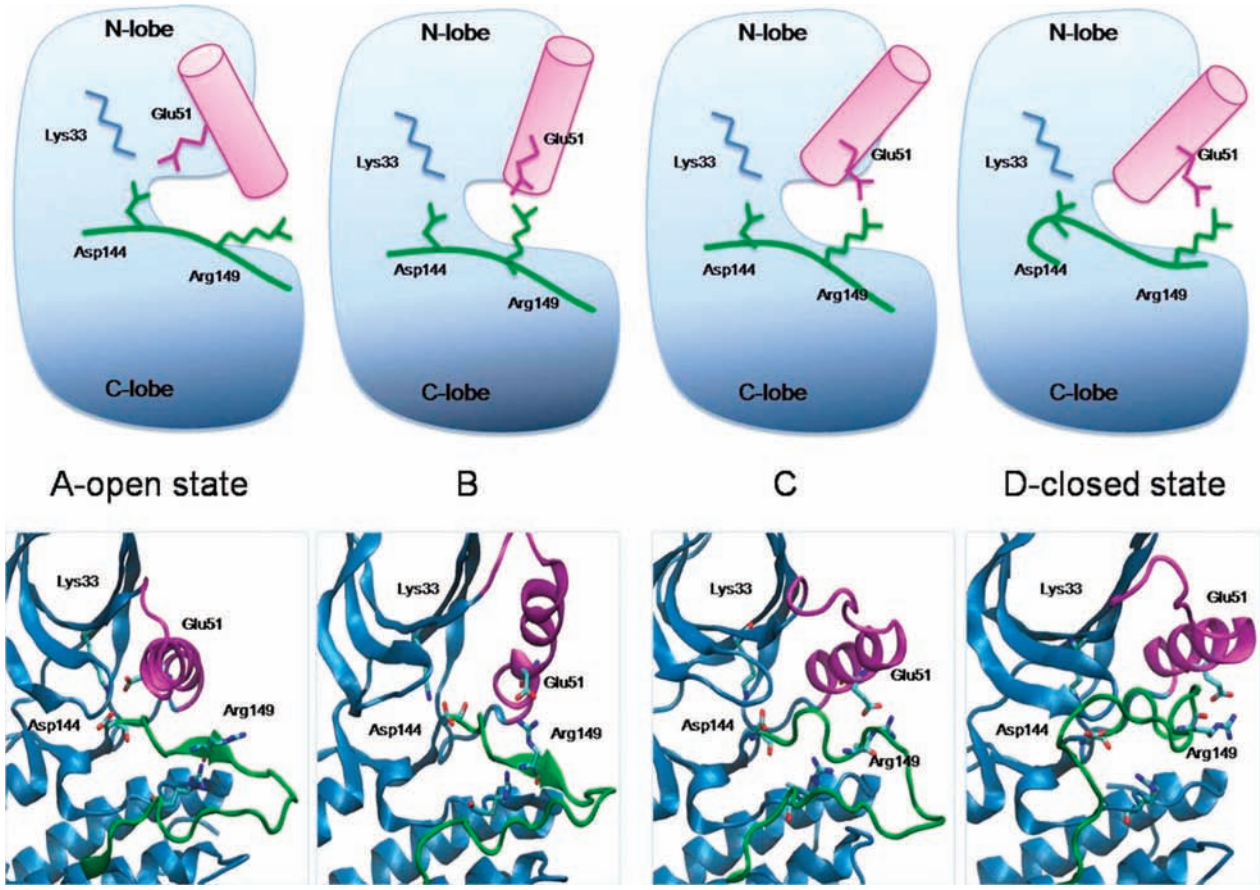

Figure 2. Different conformations assumed by CDK5 in going from the open state to the closed state: (upper panel) schematic illustration of the transition; (lower panel) corresponding cartoons of CDK5. From A to D, the figures show the progressive disruption of the Glu51−Lys33 interaction, the formation of the salt bridge between Arg149 and Glu51, the rotation of the αC-helix (pink), and the movement of the T-loop (green).

Ⓦ A movie showing the simulated open-to-closed transition in AVI format is available.

closure, not only have the αC-helix and T-loop moved, but also the G-loop is displaced, making ATP binding impossible.[60]

In the open structure, the initial segment of the T-loop forms a β-sheet motif (β9), while in the final closed structure, it is disrupted and refolds into a coil. This is at variance with the result of the homology modeling, which predicted the formation of an α-helix, but consistent with several experiments on different members of the kinase family, which failed to resolve this structural element[64,65] in the closed state.

Free-Energy Surface and Metastable Intermediate State. The use of metadynamics with the PCVs[25] also permits the reconstruction of the FES along the optimized reaction coordinate. In Figure 3, the bidimensional FES for the open-to-closed switching movement of CDK5 is reported. The FES, obtained from a 150 ns metadynamics simulation, was reconstructed as a function of the progression along the path (s) and the distance from the path itself (z) (see Methods). The projection of the FES on the s and z variables is an unusual one and probably requires a more detailed description. The z variable describes the distance from the path in the MD simulation. This has important consequences, as some energy basins that seem to be far away from the pathway are not "distinct" structures/mechanisms but indicate substantial entropic contributions to those basins. This is evident in the proximity of the closed state and is due to the intrinsic fluctuations of the activation loop. This observation is also in agreement with the findings of ref 22 that describe a very flexible activation loop in the case of the Src kinase. Additional simulations in which the overall movement was split into two subpaths were performed, and they quantitatively confirmed the FES and mechanism obtained (see SF 5 in the Supporting Information)

As expected, the open and closed states (A and D, respectively, in Figure 2) correspond to two deep free-energy basins. The closed state was 4−6 kcal/mol lower in free energy than the open one, so in the absence of an activator, the protein remains closed. The activation barrier is predicted to be ∼16−20 kcal/mol. The calculated barrier is prohibitively high for a biomolecular conformational transition. However, it was confirmed by repeated metadynamics simulations both with path-like variables and with a different set of PCVs (see SF 5 in the Supporting Information). Although we cannot rule out an error due to the force field or the existence of different pathways, the similarity of the path found here to that previously described for the homologous Src kinase, in addition to biological considerations, makes us believe that the high barrier is indeed physical. This result is consistent with a mechanism in which

(64) Cavalli, A.; Dezi, C.; Folkers, G.; Scapozza, L.; Recanatini, M. *Proteins* **2001**, *45*, 478–485.
(65) Wu, S. Y.; McNae, I.; Kontopidis, G.; McClue, S. J.; McInnes, C.; Stewart, K. J.; Wang, S.; Zheleva, D. I.; Marriage, H.; Lane, D. P.; Taylor, P.; Fischer, P. M.; Walkinshaw, M. D. *Structure* **2003**, *11*, 399–410.

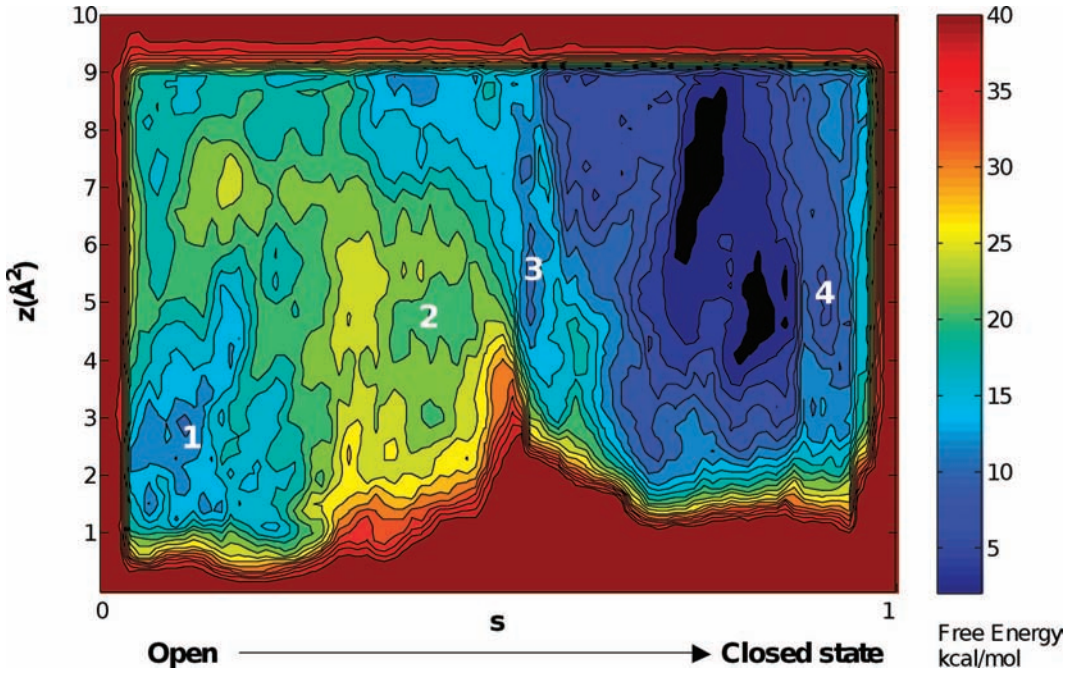

Figure 3. Free-energy surface reconstructed as a function of *s* and *z*. The energy separation between contours is 2 kcal/mol. The numbers 1−4 correspond to the states described in Figure 2.

Table 1. Distances (Å) between the Main Residues Involved in the Transition from the Open State to the Closed State

	open state	intermediate	closed state
Phe145−Leu54	5.34	9.40	13.59
Phe151−Asn121	1.95	9.39	11.73
Arg149−Leu123(C═O−H−N)	2.10	6.69	8.36
Arg149−Leu123(N−H−O═C)	2.34	5.53	6.61
Lys33−Asp144	3.39	3.46	5.28
Lys33−Glu51	3.21	14.80	16.11
Glu51−Arg149	15.32	3.86	3.95
Glu51−Arg50	11.34	9.05	6.05
Thr164−Asp126 (HRD)	3.76	4.52	1.82

p25, p35, and p39 have a fundamental catalytic role in the activation process, in agreement with previous suggestions.[59,66,67]

We report in Figure 2 the conformations corresponding to the minima observed in the FES that are most relevant in the transition from the open to the closed state (labeled as points 1, 2, 3, and 4 in Figure 3). The main geometric features of these conformations are reported in Table 1. The deepest intermediate minimum (point 3 in Figure 3) is a metastable state whose dynamical relevance was checked in an unconstrained MD run of 10 ns. During this time, the intermediate state did not exhibit any significant rearrangement. This intermediate shows structural features that differ from those of the open and closed states. The αC-helix is rotated by ∼90°, and while the Lys33−Arg149 salt bridge is formed as in the closed state, the T-loop is in an open conformation with the β9 sheet already refolded in a coil structure. In the Src and Hck kinases,[68,69] an intermediate has been crystallized in which the αC-helix is also rotated by ∼90°, but the structure otherwise differs from ours. In particular, in the Src and Hck intermediates, Glu310 in the αC-helix interacts with Arg385 and not Arg409, which corresponds to Arg149 in CDK5.

If the existence of such an intermediate were to be confirmed, it appears that enough elements of diversity exist between the intermediate and the open state of CDK5 for new drugs that target the intermediate to be designed. The existence of an intermediate state proved to be crucial in the case of Bcr-Abl tyrosine kinase,[3,70] for which it has been shown that the anticancer drug imatinib selectively binds to the intermediate.

Conclusion

The intrinsic motions of CDK5 have been investigated here by means of atomistic simulations. In 150 ns of metadynamics, we have sampled the CDK5 conformational rearrangement underlying the transition between its open and closed states. This process turned out to be rather complex, and it takes place into two steps. First, the salt bridge between Lys33 and Glu51 is broken, leading to ∼45° rotation of the αC-helix and the formation of a salt bridge between Glu51 and Arg149. Second, a highly concerted motion of the αC-helix and the T-loop leads to the closed conformation, with the αC-helix rotated by ∼90° and Arg149 and Glu51 keeping their salt bridge and being exposed to the solvent. During this step, we could detect a complete refold of the T-loop leading to the CDK5 closed conformation. In addition, we discovered the key role of Arg149 during the transition, as it worked as a linker between the N- and C-lobes, allowing a concerted movement of the two domains. The free energy of the process was accurately estimated, pointing to the closed state being more stable than the open one by 4−6 kcal/mol. Furthermore, a previously

(66) Takahashi, S.; Ohshima, T.; Cho, A.; Sreenath, T.; Iadarola, M. J.; Pant, H. C.; Kim, Y.; Nairn, A. C.; Brady, R. O.; Greengard, P.; Kulkarni, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1737–1742.
(67) Jamsa, A.; Backstrom, A.; Gustafsson, E.; Dehvari, N.; Hiller, G.; Cowburn, R. F.; Vasange, M. *Biochem. Biophys. Res. Commun.* **2006**, *345*, 324–331.
(68) Sicheri, F.; Moarefi, I.; Kuriyan, J. *Nature* **1997**, *385*, 602–609.
(69) Xu, W.; Harrison, S. C.; Eck, M. J. *Nature* **1997**, *385*, 595–602.
(70) Schindler, T.; Bornmann, W.; Pellicena, P.; Miller, W. T.; Clarkson, B.; Kuriyan, J. *Science* **2000**, *289*, 1938–1942.
(71) Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449*, 1078–1082.

undisclosed CDK5 low-energy intermediate conformation was identified and its stability assessed. A similar intermediate conformation was previously crystallized in two related protein kinases, Src and Hck,[68,69] giving support to our finding. This result, together with the similarity of the pathway found here to that described in the homologous Src kinase,[22,62,63] gives us confidence that even if what we propose is a model, its relevance might be verified by experiments such as those reported in ref 71.

From a wider perspective, this paper has reported on one of the first comprehensive atomistic investigations of kinase large-scale intrinsic motions. Remarkably, while understanding how the dynamics of kinases influences their function is a great challenge for basic science, the determination of the open-to-closed atomistic path and the associated free-energy profile may provide undisclosed kinase intermediate conformations suitable for addressing the structure-based design of potent and selective kinase inhibitors.

Acknowledgment. We thank G. Bottegoni for useful discussions at the early stage of the work. We are grateful to CSCS for providing computational resources through an advanced large-scale grant (ALPS Project). A.C. and M.R. acknowledge financial support by the Italian MUR (FIRB Project RBNE03FH5Y).

Supporting Information Available: Supporting Figures 1−7 and details concerning the additional simulations performed. This material is available free of charge via the Internet at http://pubs.acs.org.

JA806846Q