Journal of Molecular Graphics and Modelling 28 (2010) 695–706

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Why pyridine containing pyrido[2,3-d]pyrimidin-7-ones selectively inhibit CDK4 than CDK2: Insights from molecular dynamics simulation

Nahren Manuel Mascarenhas [a], Dhananjay Bhattacharyya [b], Nanda Ghoshal [a,*]

[a] Structural Biology and Bioinformatics Division, Indian Institute of Chemical Biology (CSIR), 4, Raja S.C. Mullick Road, Jadavpur, Kolkata 700 032, India
[b] Biophysics Division, Saha Institute of Nuclear Physics, 1/AF, Bidhannagar, Kolkata 700 064, India

ARTICLE INFO

Article history:
Received 4 August 2009
Received in revised form 11 December 2009
Accepted 19 January 2010
Available online 25 January 2010

Keywords:
CDK2
CDK4
Ligand binding
Molecular dynamics
Selectivity

ABSTRACT

Designing selective cyclin-dependent kinase 4 (CDK4) inhibitors is an area of intense research to develop potential anticancer drugs. The molecular basis governing the selective inhibition of CDK4 by lig17 (6-bromo-8-cyclopentyl-2-(5-piperazin-1-yl-pyridin-2-ylamino)-8*H*-pyrido[2,3-*d*]pyrimidin-7-one) has been investigated using molecular dynamics simulation. The positive charge on the ligand was determined to be an important contributor for CDK4 selectivity due to the electronegative nature of its active site. Similar studies on CDK2 indicated that Lys89 intrudes into the active site displacing the positive charge on lig17 away from the active center. This intrusion was observed to propel a drastic conformational change in lig17, weakening its binding interactions with the protein. The pyridine nitrogen (N^{AR}) of lig17 was capable of interacting with His95 (CDK4) through hydrogen bonding. N^{AR} also showed a strong tendency to mediate protein–ligand interactions through a bridged water molecule, only when bound to CDK4. The G-loop of CDK4 was observed to fluctuate extensively when complexed with lig17 and a novel "flipping-out" mechanism exhibited by $Tyr17^{CDK4/CDK4-17}$ is reported in this study. Although these proteins have similar folds, the results from principal component analysis (PCA) indicate that CDK4 and CDK2 follow an anti-correlated behavior towards the accessibility of the active site.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Cell cycle is a series of well coordinated events that result in proper cell division owing to accurate DNA replication. Two types of regulatory mechanisms have been recognized in these flow of events: a cascade of protein phosphorylations that relay a cell from one stage to the next and a set of checkpoints that monitor completion of critical events and delay/stop progression to the next stage if necessary [1]. The first line of control is primarily maintained by specific kinase family known as cyclin-dependent kinases (CDKs). CDKs are a family of heterodimeric serine/threonine kinases controlling the mammalian cell cycle. Phosphorylation by CDKs triggers the sequence of events leading to subsequent stages of cell cycle. Their activity is also tightly regulated to prevent improper cell division and unwanted cell proliferation. These CDKs require their cyclin partners and most of the CDKs require phosphorylation by CAK (cdk-activating kinase, also known as CDK7), to attain their complete activity. In mammals, different CDKs control various stages of the cell cycle and play the role of switch on/off on the cell cycle process. When on, the cell passes through the stage that the particular CDK controls, while when off, the cell cycle stops when it reaches the stage controlled by that CDK [2].

In the early stages of the cell cycle, phosphorylation of the retinoblastoma protein family, which includes pRb, p107 and p130, is scripted initially by cyclinD-dependent kinases, CDK4 and CDK6, thereby initiating the cascade of events for a successful cell cycle. Phosphorylation of pRb triggers the release of pRb regulated complexes resulting in the expression of genes required for G1 progression [3]. Due to their central role in cell cycle, several trails of evidences highlight the direct involvement of CDK4 in cancer. CDK4 and CDK6 are amplified or over-expressed in a variety of tumors and are frequent targets of genetic alterations in cancer. In addition, the natural inhibitors of CDK4/6, INK4a-ARF and INK4b loci are altered in a variety of human cancers. Also, over expression of D-type cyclins has been observed in different types of neoplasias such as breast tumors, esophageal cancer, colorectal carcinomas, head and neck carcinomas and lung cell tumors [4,5]. Further, CDK4/6–cyclinD complexes, very recently, have been demonstrated as bonafide cancer targets, especially for breast cancer [6,7].

Although several inhibitors for CDK2 [8] have been developed, only few are currently available to specifically target CDK4. Designing CDK4 selective inhibitors has been hampered by the

* Corresponding author. Tel.: +91 33 24995854/836; fax: +91 33 24735197/23967.
E-mail addresses: mailnahren@yahoo.com (N.M. Mascarenhas), dhananjay.bhattacharyya@saha.ac.in (D. Bhattacharyya), nghoshal@iicb.res.in (N. Ghoshal).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.01.008

absence of either NMR or X-ray structure. Few groups have employed computational models to successfully design CDK4 specific inhibitors [9–11]. Few theoretical studies have also been undertaken to elucidate the selective mechanism of CDK4 inhibition [12–14]. The biochemical evaluation of pyrazolopyrimidinones identified hydrogen bonding interactions to a couple of hinge residues, namely His95 and Lys22 in CDK4, to be largely responsible for the observed selectivity [11]. They coupled computational docking with synthesis and biological evaluation to put forth their proposal. A comparable study by Park et al. concluded that upon binding of selective CDK4 inhibitors the disordered loop was damped out profoundly whereas their binding in CDK2 had very little effect [12]. Studies by McInnes and group proposed that incorporation of an appropriately placed positively ionizable group in the ligand would enhance CDK4 specific inhibition due to the unfavorable positive-positive charge repulsion between the ligand and $Lys89^{CDK2}$ [10]. Noble and group with a view to establish the mechanism of selective inhibition produced CDK2 mutants bearing equivalent CDK4 residues, namely F82H, L83V, H84D, and K89T [15]. Their investigations concluded that mutation of K89T could be exploited for successfully designing CDK4 specific inhibitors for certain class of compounds, but more importantly, highlighted the significance of considering other active site residues to design selective inhibitors. In the absence of any suitable experimentally derived structures of CDK4, we have done its homology model followed by molecular dynamics simulation to understand the process of CDK4 inhibition against CDK2 binding. The crystallographic structure of CDK4 was published during the compilation of this work [16,17]. A comparative discussion on the experimental structure and the homology model of CDK4 is presented later in the manuscript.

Pyrido[2,3-*d*]pyrimidin-7-one derivatives were discovered as potent and selective inhibitors of CDK4 [18]. The ligand taken up in this study contains a positively charged center, proposed to be an essential requirement for CDK4 selectivity [10]. The authors observed a tremendous selectivity towards CDK4 upon replacing the aniline side chains to amino-pyridine containing side chains [18]. To the best of our knowledge a detailed atomistic perspective into the role of the positively charged ligand in CDK4 selectivity is yet to be investigated. It was also intended to inspect the ligand induced conformational transformations exhibited by the active site residues to accommodate the positively charged ligand in CDK2. Another significant goal of this study was to establish the role of pyridine nitrogen in ligand and its influence on CDK4 selectivity [18]. A molecular dynamics study of CDK4 versus CDK2 in presence of compound 17 (lig17 from here on) has been undertaken primarily to address the above objectives. In addition, principal component analysis (PCA) has also been performed to identify the correlated motions exhibited by these systems.

2. Computational methodology

The coordinates of CDK2 was downloaded from Protein Data Bank [19] (PDB ID: 1FIN) [20] and a homology modeled CDK4 structure, reported in our previous work [21], was used in this study. The homology model of CDK4 was constructed using the experimentally determined structures of CDK6 (PDB ID: 2EUF) [22] and CDK2 (PDB ID: 1FIN). The resulting model when evaluated using PROCHECK [23] placed 90.1% of the residues in the allowed regions of the Ramachandran plot [21] (see Supplementary Material, Fig. S1). A root-mean square deviation (RMSD) of 1.32 Å with CDK6 further indicates its high reliability. The ligand chosen for the study is based on the pyrido[2,3-d]pyrimidin-7-one core reported by Toogood and group (Fig. 1) [18]. Lig17(6-Bromo-8-cyclopentyl-2-(5-piperazin-1-yl-pyridin-2-ylamino)-8*H*-pyrido [2,3-*d*]pyrimidin-7-one) was found to exhibit excellent selectivity

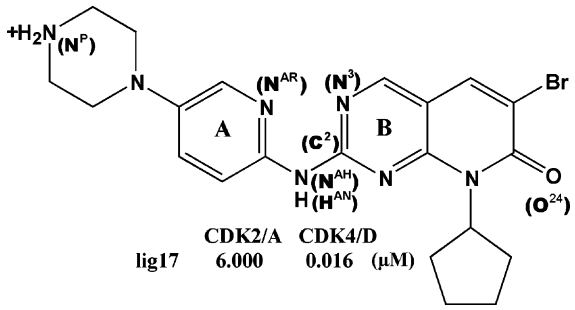

Fig. 1. The ligand selected for study. The biological activity (IC_{50}) [18] of lig17 against CDK2/cyclinA and CDK4/cyclinD are indicated in the figure. Also the atoms discussed in the texts are labeled.

towards CDK4 (0.016 μM against 6.000 μM for CDK2). The ligand was docked into the ATP pockets of CDK2 and CDK4 using GOLD (Genetic Optimization for Ligand Docking) [24]. The binding site was defined as all protein atoms within 10 Å from the carbonyl oxygen of Val96 and Leu83 in CDK4 and CDK2, respectively. The Chemscore [25] of the docked poses of lig17 in CDK2 and CDK4 were recorded as 29.085 and 33.549, respectively.

The protein as unliganded and as complex with lig17 was then subjected to molecular dynamics (MD) simulations. All simulation studies were performed using GROMACS (version 4.0.5) with GROMOS96 43a2 force field [26]. The topologies for the ligand were obtained from the PRODRG beta server [27]. The partial charges were determined by quantum chemical calculations using the WinGAMESS [28] program with RHF/6-31G* basis set employing the Mulliken population analysis using the freely available GUI WinMostar [29]. All systems were solvated in a cubical box with SPC waters keeping a minimum distance of 1 nm between the solute and each face of the box. Counter ions were added to neutralize the charge of each system. Each of the systems were then energy minimized, initially by steepest decent followed by conjugate gradient. The system was then subjected to molecular dynamics simulation using a time step of 2 fs. All bonds were restrained by applying the LINCS [30] algorithm as implemented in GROMACS. In order to allow solvent molecules to adjust to the solute, a small MD run of 40 ps was performed, restraining the protein atoms. The systems were then relaxed from restrains and 12 ns of MD simulation were initiated. The first 1 ns of initial run was considered to be needed for equilibration and hence the next 11 ns of run was only included for analysis. All systems were run under periodic boundary conditions and the coordinates of MD simulations were saved every 500 steps (1 ps). The non-bonded interactions were truncated at 12 Å with long range electrostatic interactions modeled using PME beyond the cut-off with the neighbor list updated every 10 steps. Temperature was coupled to a reference bath at 300 K using Berendensen thermostat [31] while the pressure was maintained at 1 bar using the Parrinello-Rahman scheme [32]. The protein, solvent and ligand were independently coupled to the temperature bath. Analyses of the simulations were performed using analysis tools available in GROMACS while the trajectories were visualized with VMD [33].

The interaction energies were calculated based on the non-bonded interactions, electrostatic (E_{Coul}) and van der Waals (E_{LJ}) modeled as Coulombic and Lennard-Jones parameters, respectively, between the ligand and the protein. Electrostatic potential was calculated using Delphi [34] module in InsightII. Principal component analysis (PCA) [35] was performed using analysis tools of GROMACS package, considering only the backbone atoms for generating the covariance matrix. Prior to performing PCA, all translations and rotational motions were eliminated by fitting the trajectory to a reference structure, in this case the average structure. In order to visualize the motions, the distance between

the center of mass (COM) of highly fluctuating segments between the first and the last frame of the projection of the corresponding eigenvector was calculated by means of Tcl scripts in VMD.

3. Results and discussion

The attempt to place lig17 in the active site of CDK2 by docking produced very similar binding pose as those observed earlier in our docking results with CDK4 [21] as well as similar to the CDK6-ligand complex deposited in PDB (PDB ID: 2EUF) [22]. In CDK2-17, the amino group extending from the pyrido[2,3-d]pyrimidin-7-one core donates a hydrogen bond to Leu83 while the N^3 atom of the core accepts a hydrogen bond from the backbone amide of Leu83. The core is highly stabilized in a hydrophobic environment formed by residues Ile10, Ala31, Val64, Phe80, Leu134 and Ala144. The positively ionizable piperazine ring is exposed to the solvent accessible side of the ATP pocket. In CDK4-17 complex, Val96 (equivalent to $Leu83^{CDK2}$) is involved in hydrogen bonding interaction with lig17. The only notable difference observed among the docked poses of lig17 in CDK4 and CDK2 is the presence of a hydrogen bond involving the $His95^{CDK4}$ and the pyridine nitrogen (N^{AR}) of the ligand. The simulation study should in principle evaluate the stability of each of these interactions illuminating the underlying mechanism of selective CDK4 inhibition.

A well established strategy to monitor the attainment of equilibration involves the investigation of root mean square deviation (RMSD) of atoms from their initial positions (see Supplementary Material Fig. S2). The figure reports RMSD of Cα atoms after least square fitting the corresponding Cα from different frames to the starting structures of the simulation. Both ligand bound and unliganded form of CDK2 as well as CDK4 seem to exhibit a similar conformational freedom from their starting structures. The plot also indicates that it takes about 1 ns for the system to equilibrate and hence the first 1 ns of the simulation was not used for analysis to minimize equilibration artifacts.

In order to evaluate regions exhibiting significant dynamic behavior, the root mean square fluctuations (RMSF) of the Cα atoms were calculated (Fig. 2). The fluctuations correspond well with loop regions, which are devoid of a defined secondary structure and expected to be labile. The G-loop in both CDK2 and CDK2-17 shows identical behavior while it behaves differently in CDK4. When complexed with lig17, the fluctuations of G-loop in CDK4 seem to get amplified. Thus complex formation seems to impart more mobility to the G-loop in case of CDK4 than in CDK2. The second region of significance is the disordered loop that lies between β3-sheet and α1-helix. The behavior of disordered loop is quite similar in both CDK4 and CDK2, and in both cases, ligand binding seems to induce additional flexibility. Our observation contradicts the conclusion drawn by Park et al. [12], that the presence of selective inhibitor restricts the fluctuations of the disordered loops extensively in CDK4 than in CDK2. In our case the disordered loop of CDK4 was found to be stable in the unliganded form but fluctuated extensively upon ligand binding. It is to be noted that their conclusion was based upon a simulation runtime, which is comparatively much smaller. Since disordered loops are highly fluctuating segments, their fluctuations are anticipated to get accumulated over time resulting in high RMSF values. The α2-helix (Fig. 2, k2/k4) that immediately follows the hinge region displays restricted motions in complex state in both systems indicating the impact of ligand binding and its role in stabilizing this helix. In case of CDK4, ligand's presence seems to induce more stability to almost whole of the C-terminal lobe while such stabilizing effect is not observed in CDK2. The T-loop residues of CDK2 become more labile when bound to lig17 than in unliganded form. In contrast, the corresponding regions in CDK4 and CDK4–17 do not exhibit such a behavior. Residues 235–265 that show high amplitude fluctuations in unliganded CDK4, exhibit restricted motion when complexed with lig17.

The final residues constituting the α1-helix in CDK2-17 were alone observed to display enhanced fluctuations. Such perturbations to α1-helix have been documented in an earlier study [36], but the extent of α1-helix misfold was much less in our case. Analysis of secondary structure elements using DSSP [37] confirms that only in CDK2-17, the final residues of the α1-helix misfold during the last 5 ns of the simulation (see Supplementary Material Fig. S3). The β1-sheet of CDK2 is disrupted initially but gains its shape and maintains a constant structure in the last 9 ns of the simulation. In case of CDK2–17, β1 structure attains stability only in the last 6 ns of the simulation. In comparison, the β1-sheet is unstable in CDK4, more so in CDK4-17. In case of CDK4, β1 is preserved at least up to 7 ns (approximately) of the simulation and only then it seems to disappear. On the other hand, the β1-sheet is itself not exclusively constituted in CDK4-17. G-loop was found to exhibit extensive fluctuation only in CDK4-17 while such widespread perturbations were not seen in others. The G-loop is flanked by two beta-sheets namely, β1 and β2. Since β1-sheet is itself not exclusively formed in CDK4-17, it could implant additional mobility into the following G-loop. Thus it is proposed that the high amplitude fluctuations exhibited by G-loop of CDK4-17 could be a direct outcome of its inability to form the preceding β1-sheet. Except these minor differences, all four structures appear to comprise similar structural folds throughout the duration of simulation.

To express the strength of ligand affinity towards the two proteins, the interaction energy of lig17 with CDK2 and CDK4 were

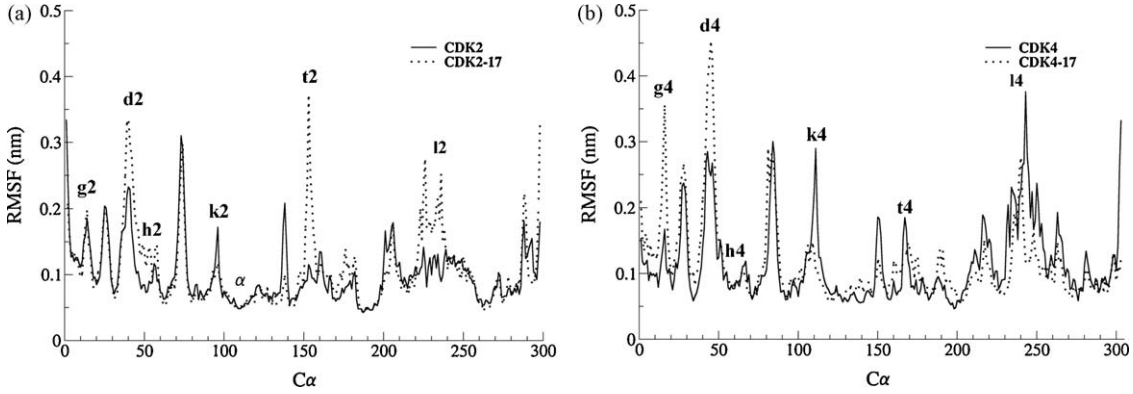

Fig. 2. RMSF values of the Cα (alpha) atoms in (a) CDK2 and CDK2-17 and (b) CDK4 and CDK4-17 during the last 6 ns of the simulation. Regions labeled correspond to different segments of the protein (2-CDK2 and 4-CDK4). (i) G-loop, g2 (CDK2, 10–18) and g4 (CDK4, 12–20); (ii) disordered loop, d2 (37–44) and d4 (39–49); (iii) α1-helix, h2 (45–58) and h4 (50–67); (iv) α2-helix, k2 (87–93) and k4 (100–106); (iv) T-loop, t2 (147–165) and t4 (160–177); (v) L14 loop, l2 (220–247) and l4 (231–256).

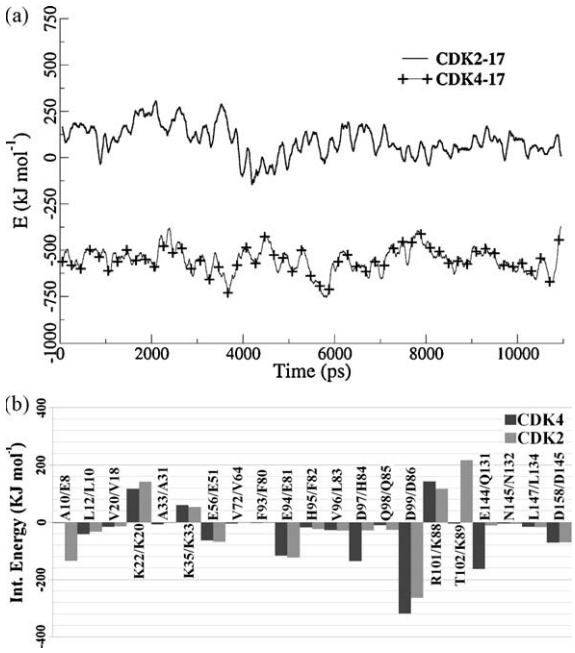

Fig. 3. (a) The total non-bonded protein–ligand interactions calculated over the trajectory frames. The plot values are original values averaged over 100 data points (i.e. 0.1 ps) to gain better comparisons. (b) Interaction energy of lig17 with individual amino acid residues forming the active site in CDK2 and CDK4.

measured (Fig. 3(a)). Throughout the simulation lig17 interacted efficiently with CDK4 (avg. interaction energy -743.328 ± 107.934) than with CDK2 (-106.553 ± 115.468) (see Supplementary Material, Fig. S4). In order to get a better insight into the contribution of individual amino acids towards ligand binding, the interaction energies of amino acids surrounding the active site were evaluated. The primary locus of difference in the interaction patterns of lig17 with CDK4 and that with CDK2 arises from the fact that the non-favorable contribution of $Lys89^{CDK2}$ is transformed into a favorable contribution by the occupation of $Thr102^{CDK4}$. Also the negatively charged $Asp99^{CDK4}$ interacts better with lig17 than the corresponding $Asp86^{CDK2}$ and the replacement of $His84^{CDK2}$ by $Asp97^{CDK4}$ further boosts its electrostatic interaction with ligand (Fig. 3(b)). However, the positively charged lig17 interacts with $Glu8^{CDK2}$ extensively while its replacement by $Ala10^{CDK4}$ reduces this interaction quite considerably.

The evaluation of the vdW component of total interaction energy in CDK2-17 (avg. -198.448 ± 15.563) and CDK4-17 (avg. -189.63 ± 13.744) as well as the electrostatic component in CDK2-17 (91.896 ± 112.819) and CDK4-17(-553.697 ± 106.967) confirms a strong Coulombic attraction operates only in the latter. McInnes and group [10] had earlier shown that the electronegative nature of CDK4 active site in comparison to CDK2 is due to the replacement of $Lys89^{CDK2}$ by $Thr102^{CDK4}$. To rationalize such a strong preference of lig17 towards CDK4, the electrostatic potential of the two proteins were compared. The electrostatic potential was calculated on the final structures stored during the MD run and it indicated that the entrance of the ATP pocket is negatively and positively charged in CDK4-17 and CDK2-17, respectively (Fig. 4). Asp99 and Thr102 flank the entrance of the CDK4 active site, while Asp86 and Lys89 take the corresponding position in CDK2, leading to the difference in electrostatic potential. The side chains of $Asp99^{CDK4}$ and $Thr102^{CDK4}$ were also observed to exhibit hydrogen bonding interaction for about 81% duration of the simulation, holding the two residues in close proximity and leading to the negative characteristic of its active site. From the figure it could be acknowledged that in CDK4-17 complex the piperazine part of the ligand is oriented towards the solvent accessible part of the ATP pocket, bending away from the hinge, while it is oriented towards C-terminal (Leu298) carboxyl group in CDK2-17. This difference in orientation of lig17 is a direct outcome of the intrusion of Lys89 at the

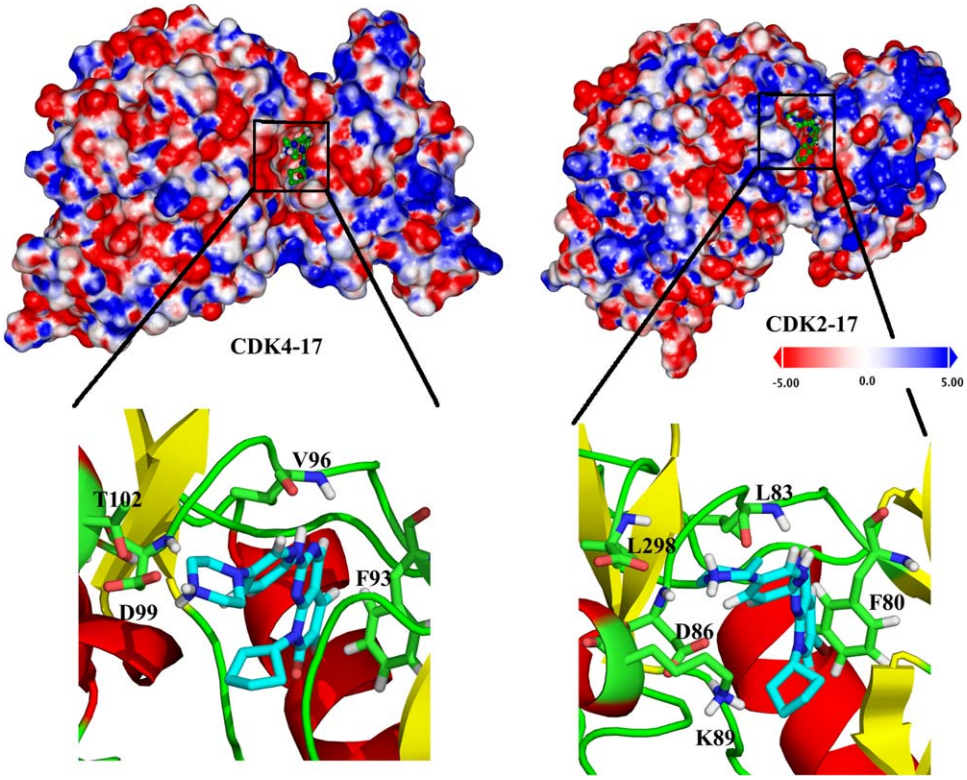

Fig. 4. The electrostatic potential (kT/e) mapped on the surface of the protein using program Delphi on CDK4-17 (a) and CDK2-17 (b) on the last snapshot of the MD run. (c) and (d) Projects a close-up view of the active site of the respective proteins.

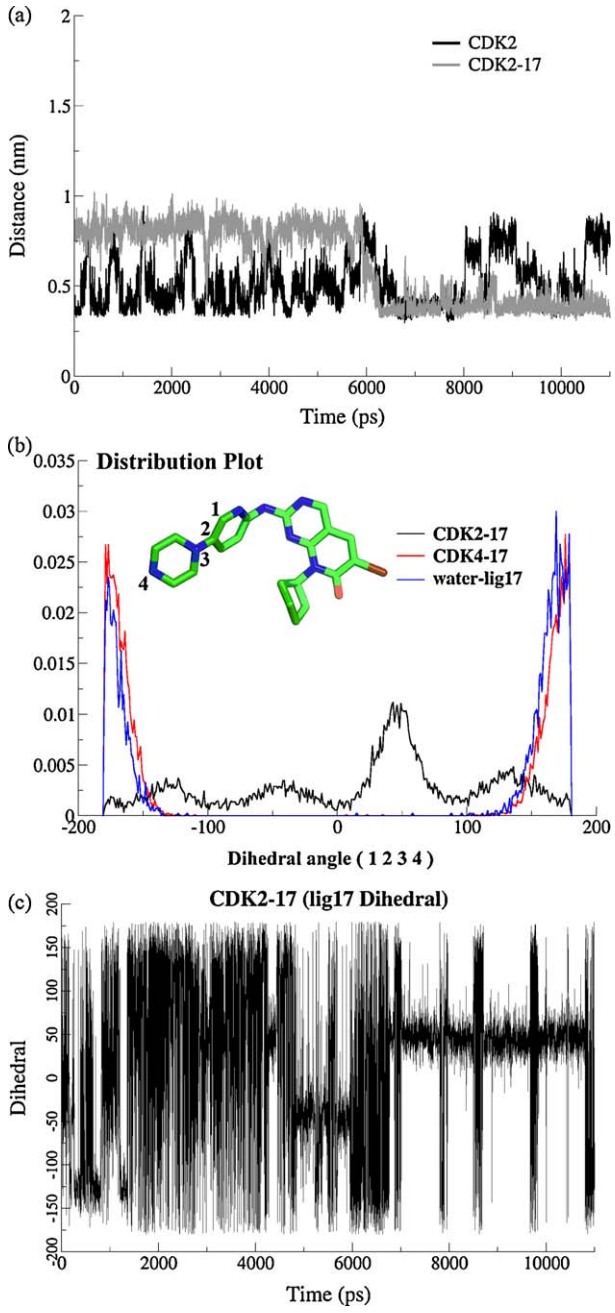

Fig. 5. (a) The distance of Asp86(CG). . .Lys89(NZ) plotted as a function of time in CDK2-17. (b) Distribution plot of the psedo-dihedral angle formed by 1_2_3_4 in CDK2-17, CDK4-17 and lig17-water simulation. (c) Variation in the pseudo-dihedral (1_2_3_4) of lig17 plotted as a function of time in CDK2-17.

active site. The Asp86 (CG). . .Lys89 (NZ) distance, plotted as function of time explains the behavior of Lys89 in CDK2 (Fig. 5(a)). In case of unliganded CDK2, the distance between the two atoms fluctuates between 0.3 and 0.9 nm, but in complex with lig17, the distance between the two atoms exists in two states. During the first 5 ns the value fluctuates around 0.8 nm and it falls to 0.35 nm in the last 5 ns, highlighting the strong electrostatic forces operating between the two residues in the later part of the simulation. The strong electrostatic pull exerted by Asp86 enables Lys89 to intrude the active site of CDK2 pushing the piperazine side chain of lig17 upward (Fig. 6).

The positive charge on Lys89 at the entrance of the active site should in principle create a significant stress in CDK2-17 complex. As a result it is natural to expect a conformational tweak of lig17$^{CDK2-17}$ in such a way to minimize these strong electrostatic repulsions. To assess this conformational switching mechanism the pseudo-dihedral angle formed by atoms 1, 2, 3 and 4 in lig17 were measured in CDK2-17, CDK4-17 and lig17-water (performed for 3 ns, data not shown) simulations (Fig. 5(b)). The distribution plot of the pseudo-dihedral angle confirms that the conformation of lig17-water is quite identical to that of CDK4-17 complex but differs substantially with that of CDK2-17. In CDK2-17, a range of conformations is explored by lig17, which directly reflects its instability and inability to emulate strong interactions with the protein. Also, in the last 5 ns of the simulation, lig17$^{CDK2-17}$ is transformed to adopt a conformation that shifts the positive charge on the piperazine a little upward towards the α2-helix. This results in the dihedral angle to be maintained predominantly at around 50° during the final stages of simulation (Fig. 5(c)). Although this conformational switching enhances the possibility of its interaction with carboxyl terminal of Leu298 (Fig. 6), it certainly becomes very different from the conformation observed for lig17 in free simulation (see Supplementary Material, Fig. S5) and indicates that it is transformed to an unstable conformation. On the other hand, in CDK4-17 complex, the lig17 conformation resembles closely the conformation observed in lig17-water simulation (Fig. 5(b)). As no drastic conformational changes were observed for lig17 in CDK4-17, it suggests that lig17 is more stable as a complex with CDK4 than with CDK2.

The crystallographic structures of CDK4 [17] were made available during the compilation of the work and hence it was considered imperative to discuss the difference between the experimental structure and our theoretical model. To compare the folding patterns, the secondary structural elements of the two structures were compared following the Kabsch-Sander [37] definition (see Supplementary material, Fig. S6). It is found that both the structures contain similar secondary structural folds, which further validate the homology model used in the study. A RMSD of 4.76 Å for the Cα atoms between the model CDK4 structure and that of experimental structure CDK4/CyclinD3 (PDB ID: 3G33) [17] confirms a huge conformation difference between the two proteins, most of which can be due to inter-domain movements. Based on the crystallographic evidence it was established that the experimental structure of CDK4/cyclinD3 resembles the structures of inactive and cyclin-unbound CDK2 [38] and CDK7 [39] as well as inactivated INK-inhibited structure of CDK6 [40]. Since our homology model of CDK4 is based on the active conformation of CDK6 and CDK2, the high difference in the RMSD values is not surprising. Further, a RMSD value of 3.92 between the cyclin-unbound CDK2 [41] (PDB ID: 1HCK) and CDK2/cyclinA complex [20] (PDB ID: 1FIN) justifies the huge conformational difference induced upon cyclin binding. On the other hand, a RMSD of 1.32 Å between the homology modeled CDK4 and the activated CDK6 confirms that the structure resembles to that of an activated CDK conformer.

There are two major differences between the cyclin-bound and cyclin-unbound structures of CDK2 [42]. First, the T-loop in cyclin-unbound CDK2 emerges from the C-terminal domain to block the entrance for the substrates to access the catalytic site. In comparison, the T-loop in CDK2/cyclinA moves back into the C-terminal domain and paves way for the substrate to access the catalytic cleft. In complete contrast, CDK4 fails to shift the T-loop into the C-terminal domain even in the presence of cyclinD3. The T-loop exhibits a short helix (residues 162–171) and remains obstructing the way for the substrate to access the catalytic cleft [17] (see Supplementary Material Fig. S7). The second major difference is centered around the α1-helix, conserved across all family of CDKs. Glu51^{CDK2} (Glu56^{CDK4}) of the helix, critical for catalysis, in the cyclin-unbound form is oriented away from the catalytic site and remains exposed to the solvent. Upon cyclinA

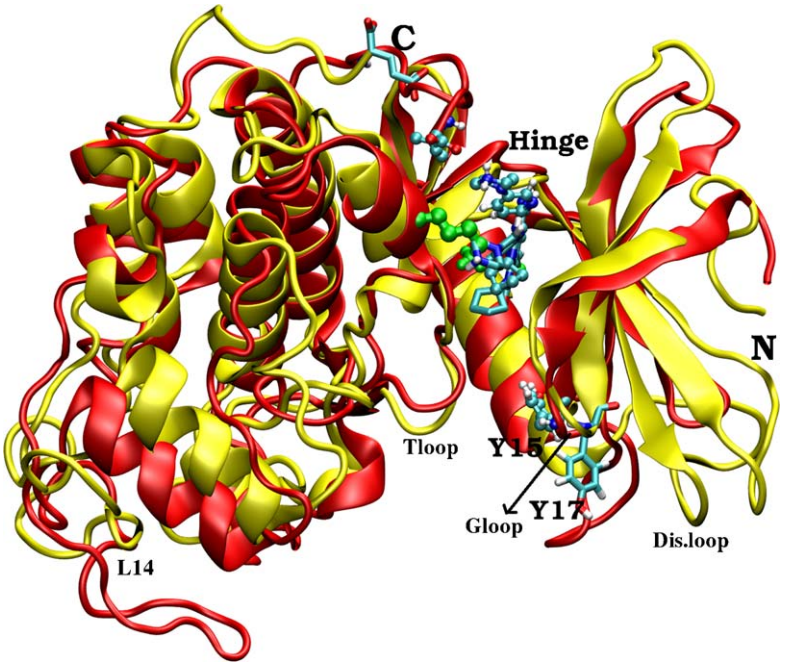

Fig. 6. Structural alignment of CDK2-17 (red) and CDK4-17 (yellow) performed on the final structure stored during MD. Tyr15, Leu298 and lig17^{CDK2} is shown in ball and stick while the corresponding residues in CDK4 Tyr17, Glu303 and lig17^{CDK4} are shown in cylindrical representation. Also displayed in the figure is the intrusion of Lys89^{CDK2} (green) into the active site.

binding and its extensive interaction with α1-helix pushes Glu51^{CDK2} into the catalytic region. CyclinD3, on the other hand, does not extensively interact with CDK4 and is unable to shift Glu56^{CDK4} into the active site. Hence, the crystallographic structures reveal that cyclinD3 does not induce the observed conformational transformation into CDK4 associated with CDK activation upon cyclin binding. So the structure of CDK4 resembles closely the structure of an inactive CDK, making it unsuitable for studying protein–ligand interactions. On the contrary the homology modeled structure resembles that of an activated CDK conformer. Its T-loop is towards the C-terminal lobe and also the Glu56 in α1-helix is oriented towards the active site to form the catalytic triad along with Lys35 and Asp158. These features put together essentially highlight that the homology model of CDK4 resembles the activated form of CDKs making it better suited for the objective under consideration.

In several of the CDK2-inhibitor complexes, inhibitors imitate the hydrogen bonding pattern of adenine in ATP. The hinge region of CDK2 (residues 81–84) possesses a set of hydrogen bond donor and acceptor sites which several of the inhibitors target [43]. In CDK2 complex with ATP, the adenine ring forms a couple of hydrogen bonds with hinge residues Leu83 and Glu81. N1 of ATP accepts a hydrogen bond from the backbone nitrogen of Leu83 while N6 donates a hydrogen bond to the carbonyl oxygen of Glu81 [20]. Hence, an inhibitor with its ability to maximize the hydrogen bonding capacity to these residues would certainly possess a greater affinity towards the active site. Further evidence in this direction is provided in CDK6/CyclinV-PD033291 complex [22], where the ligand PD033291 is an analogue of lig17. In this co-crystalline complex, Val101 amide (equivalent to Leu83 and Val96 in CDK2 and CDK4, respectively) donates a hydrogen bond to the N01 (equivalent to N^3 in lig17) nitrogen from pyrido[2,3-d]pyrimidin-7-one core while the Val101 backbone carbonyl oxygen accepts a hydrogen bond from the HN04 (equivalent to H^{AN} in lig17) of ligand. Hence, monitoring the hydrogen bonding ability of lig17 with CDK4 and CDK2 could give direct clues about its affinity towards the respective proteins. As the first step, all possible hydrogen bonds in CDK4-17 and CDK2-17 were monitored as a function of time (Fig. 7, Table 1). Hydrogen bonding interactions are more predominant in CDK4-17 complex than in

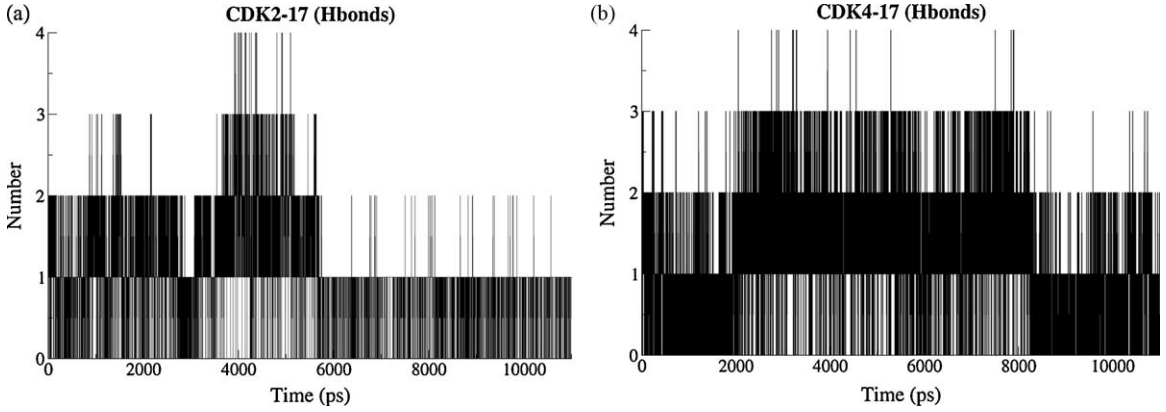

Fig. 7. The total number of intermolecular hydrogen bonding interaction exhibited by lig17 in complex with (a) CDK2 and (b) CDK4 as a function of time.

Table 1
Protein–ligand hydrogen bonding interaction pairs extracted from the trajectory.

CDK217			CDK417		
Donor	Acceptor	Occupancy[a]	Donor	Acceptor	Occupancy
LEU83-Main	Lig17	69.67	VAL96-Main	Lig17	51.04
Lig17	LEU83-Main	29.54	Lig17	VAL96-Main	42.07
LYS33-Side	Lig17	7.01	HIS95-Side	Lig17	13.70
Lig17	LEU298-Main	2.36	Lig17	THR102-Side	5.31
Lig17	ARG297-Main	0.44	LYS35-Side	Lig17	4.26
Lig17	ASP86-Side	0.39	Lig17	HIS95-Side	0.81
Lig17	GLN85-Side	0.35	Lig17	ASP99-Side	0.63
GLN85-Side	Lig17	0.12	GLY13-Main	Lig17	0.09
ASP145-Main	Lig17	0.01	Lig17	ASP97-Main	0.05
LYS89-Side	Lig17	0.01			
last 5 ns			last 5 ns		
Donor	Acceptor	Occupancy	Donor	Acceptor	Occupancy
LEU83-Main	Lig17	83.9	Lig17	VAL96-Main	46.69
Lig17	LEU83-Main	0.60	VAL96-Main	Lig17	38.79
Lig17	ARG297-Main	0.46	HIS95-Side	Lig17	9.74
LYS89-Side	Lig17	0.02	LYS35-Side	Lig17	7.26
			Lig17	THR102-Side	1.98
			Lig17	ASP99-Side	0.74
			Lig17	HIS95-Side	0.30
			GLY13-Main	Lig17	0.10
			Lig17	ASP97-Main	0.04

[a] Occupancy of hbonds as % of the simulation time.

CDK2-17. To highlight in quantitative terms, we measured the percentage occupancy of all possible hydrogen bonding pairs. As expected, Leu83^{CDK2} and Val96^{CDK4} contribute extensively towards hydrogen bonding interaction. Excluding these residues, His95, Thr102 and Lys35 were also capable of forming hydrogen bonding interaction with lig17 in CDK4 but only for 14%, 5% and 4% of the duration of the simulation time, respectively. In CDK2-17, other than Leu83^{CDK2} only Lys33 and Leu298 were able to form considerable amount of hydrogen bonding interactions with the ligand. Except Leu83(donor)–lig17(acceptor), the hydrogen bonding ability of lig17 with CDK2 is almost lost during the last 5 ns of the simulation, while in CDK4-17 complex all the hydrogen bonding interactions were still quite active (Table 1). The last 5 ns of the simulation were earlier shown to correspond with the time when Lys89 intrudes into the entrance of the active site compelling lig17 to adopt a very unstable conformation. Thus this intrusion also seems to have a direct impact on hampering the hydrogen bonding ability of lig17. The Leu83(CO)-H^{AN}(lig17) hydrogen bonding was observed to be lost in the last 5 ns of the simulation in CDK2-17 complex while the corresponding interaction in CDK4-17 between Val96(CO)-H^{AN}(lig17) was observed for 47% of the duration of the simulation. A residence time of 84% and 39% for the hydrogen bonding interaction Leu83(NH)-N^{AH}(lig17) and Val96(NH)-N^{AH}(lig17) respectively, suggests that the hydrogen bonding interaction involving the hinge donor is quite stable but that involving Leu83(CO)^{CDK2} and Val96(CO)^{CDK4} is reasonably different. This can particularly direct lig17 to be more selective towards CDK4.

The G-G–G segment of the G-loop is highly conserved across the CDK family. Comparison of G-loop sequences in CDK2 (^{10}GEGTYGVV^{18}) and CDK4 (^{12}GVGAYGTV^{20}) indicates two important changes. The replacement of Glu12^{CDK2} and Thr14^{CDK2} by Val14^{CDK4} and Ala16^{CDK4} makes G-loop little more hydrophobic in CDK4. In the light of this difference, the solvent accessible surface area (SASA) of G-loop was measured to account for the behavior of G-loop. Upon ligand binding the G-loops in both CDK2-17 and CDK4-17 get better exposed to the solvent than their unliganded counterparts (see Supplementary Material Fig. S8). The RMSF analysis indicated G-loop to fluctuate predominantly when complexed with lig17 (Fig. 2). To investigate and further understand about the fluctuating behavior of G-loop, it is compared with that of CDK6 (^{19}GEGAYGKV^{27}). It was not possible to compare the G-loop from MD simulations to that in the crystallographic structure of CDK4 due to its inactive nature as discussed earlier. In several of the CDK6 structures, deposited in protein data bank (PDB), the G-loop regions were either completely missed or were located with high temperature factors. For example, in the co-crystalline complex of CDK6/CyclinV-aminopurvalanol [PDB ID: 2F2C] residues 21–24 were not established with other regions of the G-loop located with high temperature factors. In co-crystalline complex of CDK6/CyclinV-PD033291, [PDB ID: 2EUF], although backbone of the G-loop were ascertained, the side chains of Glu21, Tyr24 and Lys26 were not located [22]. In a similar complex structure of CDK6/CyclinV-fisetin [PDB ID: 1XO2], again several side chains in the G-loop were not estimated [44]. Thus the crystallographic structures of CDK6 strongly communicate that its G-loop is highly fluctuating. On the other hand, CDK2 complexes do not possess very high temperature factors at the G-loop. Since CDK6 is highly homologous to CDK4, and from the analyses of deposited inhibitor-CDK6 complexes, a similar G-loop behavior of the latter when complexed with lig17 could not be underrated. Thus the high mobility associated with G-loop of CDK4-17 bears close resemblance to inhibitors complexed with CDK6 while remaining appreciably different from CDK2.

Another unique feature extracted from the simulation was concentrated at the Tyr17^{CDK4} (equivalent to Tyr15^{CDK2}). In comparison to two residues (Thr14 and Tyr15) in CDK2, only a single phosphorylable residue (Tyr17) is present in CDK4. Phosphorylation of Tyr15 in CDK2 has been found to enhance G-loop's flexibility. It was proposed that phosphorylated-Thr15 residue severely restricts substrate binding as well as its orientation in the active site without considerably altering the overall structural topology of CDK2 [45]. Earlier, modelling studies had also proposed the increased mobility installed to G-loop upon phsophorylation of Tyr15 [46]. To gain a comparison of the orientation of Tyr15^{CDK2} and Tyr17^{CDK4} the Chi1 (χ^1) dihedral angle was measured. The distribution plot indicates that the occupancy of Tyr15 is quite opposite to that of Tyr17. In both CDK4 and CDK4-17, Tyr17 adopts gauche+ conformation where as in CDK2 and CDK2-17 it prefers to remain in gauche− (see

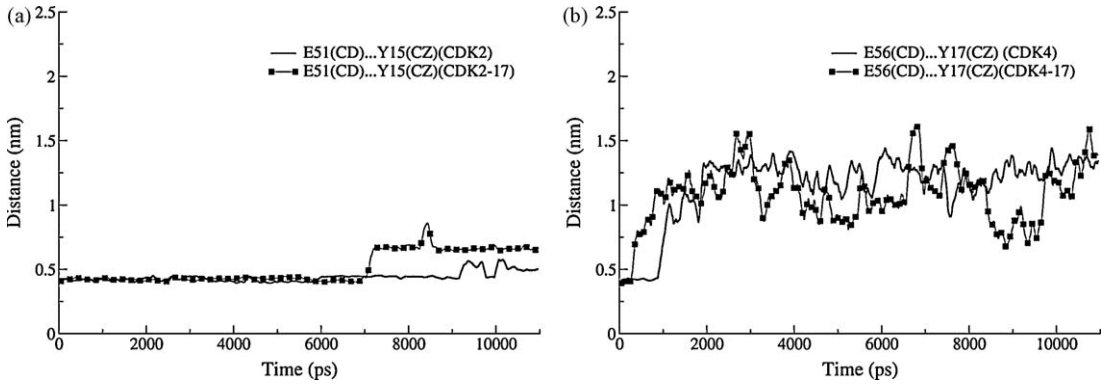

Fig. 8. The distances (a) Glu51(CD)-Tyr15(CZ)^{CDK2} and (b) Glu56(CD)-Tyr17(CZ)^{CDK4} plotted as a function of time to highlight the difference in orientation of the conserved tyrosine in the G-loop.

Supplementary Material, Fig. S9). In both ATP-bound CDK2 [21] as well as inhibitor bound CDK2 [47], the Tyr15 Chi1 angle was found to be 75.99° and 71.71° respectively, and was also oriented towards the Lys33-Glu51-Asp145 triad. It is apparent from these values that the orientation of Tyr15$^{CDK2/CDK2\text{-}17}$ observed in our simulation is quite similar to what has been observed experimentally. From these values it is noticeable that gauche− conformation observed in CDK2 are also well preserved in our simulation. Additionally a Chi1 value of −78.46° for the phosphorylated Tyr15^{CDK2} [PDB ID: 2CJM] gives the first indication that phosphorylated Tyr15^{CDK2} is quite similar to Tyr17$^{CDK4/CDK4\text{-}17}$. Another approach to assess this characteristic difference in the orientation of Tyr17^{CDK4} and Tyr15^{CDK2} is to measure its distance from Glu56^{CDK4} and Glu51^{CDK2} respectively. Since glutamate is present at the α1-helix, the analysis would clearly highlight if the residues were solvent exposed or buried. The plot indicates that in both CDK2 and CDK2-17, Tyr15 is very close to Glu51 while Tyr17 lies far away from the corresponding Glu56 in CDK4 and CDK4-17 (Figs. 8(b) and 6). Also Tyr15 in both CDK2 and CDK2-17 was observed to exhibit significant hydrogen bonding interaction with Glu51^{CDK2} while Tyr17^{CDK4} shows no significant evidence of such an interaction. Tyr15^{CDK2} was found to be involved in hydrogen bonding interaction with Glu51 and Asp145 for about 73% and 12%, of the duration of the simulation, respectively. Upon lig17 binding it exhibited intermolecular hydrogen bonding with Glu51 and Thr47 for 53% and 20% of the duration of the simulation, respectively. In both CDK4 and CDK4-17, Tyr17 did not show any appreciable (>10%) hydrogen bonding interaction with any residues of the protein. Tyr17 and Glu51 exhibited strong hydrogen bonding affinity only in the initial stages, during the first 500 ps of the simulation and thereafter it seems to completely disappear. The above results further accumulate evidence in support of Tyr17 being solvent exposed in CDK4 and in its complex with lig17. Water-pi (O-H…pi) interactions involving hydrogen atoms of water and pi electron cloud (aromatic rings) have been investigated [48]. Recently, Sankararamakrishnan and group have reported the possibility of even lone pair···pi interactions between water oxygens and aromatic residues [49]. It is proposed that the flipped Tyr17$^{CDK4/CDK4\text{-}17}$ could possibly attain stability by such type of interactions.

A remarkable selectivity shift towards CDK4 was observed upon substitution of aniline with 2-aminopyridine [18]. The pyridine nitrogen (N^{AR}) was thus adjudged to play a pivotal role in shifting the selectivity of the ligand confirming that it should mediate certain interactions in CDK4 not prevalent in CDK2. In view of this observation, an effort to identify specific interaction involving N^{AR} and protein was initiated. In the docking studies, a hydrogen bond involving N^{AR} and His95^{NE2} was observed [21]. It prompted us to analyze the trajectory and check for the existence of this hydrogen bonding interaction during the course of the simulation. Monitoring the distance between the two atoms confirm that the favorable distance for hydrogen bonding is found only during 2–8 ns of the simulation time (see Supplementary Material Fig. S10). In this period, the distance between these atoms are optimal (<0.35 nm) for about 80% of the simulation time to interact by hydrogen bonding. A cutoff of 30°, 45° and 60° estimates 24%, 47% and 64% possibility of hydrogen bonding, respectively, between N^{AR} and His95^{CDK4}. A similar kind of interaction was not possible in CDK2-17 complex due to the presence of Phe82^{CDK2} as the corresponding residue.

The lack of any interaction involving N^{AR} in the remaining 5 ns of simulation certainly does not advocate its role in selectivity. Analysis of the trajectory led us to identify water molecules clustered around N^{AR} in CDK4-17 complex, speculating the possibility of protein-ligand interaction through a water bridge. The involvement of water mediated interaction in stabilizing protein–ligand complex is a well established phenomenon. Theoretical studies have also illuminated the necessity to probe such interactions to gain a better understanding of the inhibitory activity [50]. Kriz et al. have identified water molecules that arrange interaction between protein-ligand in few well known inhibitors of CDK2 [51]. Further, a recent CDK2-inhibitor co-crystalline structural complex also highlights the importance of water mediated interaction [52]. These experiments clearly signify water centered interactions are an integral part of protein-ligand interactions in these systems. Hence, the investigation was extended to mine the possibility of water-mediated protein-ligand interactions through N^{AR} (Fig. 9, Table 2). For investigating water molecule that organizes interaction between the protein and ligand through N^{AR}, the whole of the protein was selected followed

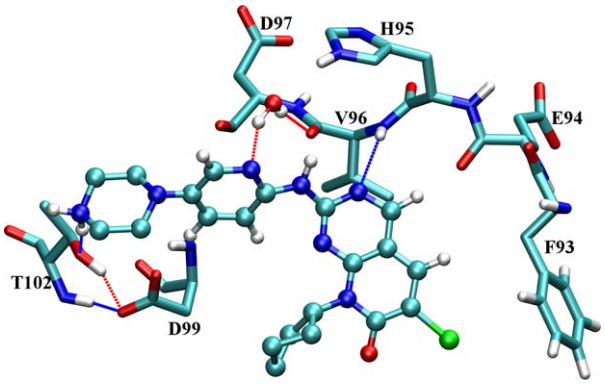

Fig. 9. A snapshot demonstrating one of the possible water-mediated interaction between the protein and ligand through pyridine nitrogen (N^{AR}). Other notable hydrogen bonding interactions are also revealed in the figure.

Table 2
Water-mediated interaction mediated by N^{AR} and O^{24} of lig17 with CDK4 and CDK2.

	1–2000	2001–4000	4001–6000	6001–8000	8001–11000
CDK417					
AC-w[a]-N^{ARb}	1128	139	90	142	1497
Hinge(93-99)-w-N^{AR}	1109	115	63	122	1434
AC-w-O^{24}	1666	1148	1692	1248	1558
Asp158-w-O^{24}	990	289	198	175	157
Glu56-w-O^{24}	0	0	0	4	0
Lys35-w-O^{24}	0	0	0	1	2
Glu144-w-O^{24}	968	679	1323	566	644
Asn145-w-O^{24}	88	11	285	623	713
CDK217					
AC-w-N^{AR}	630	656	802	156	269
Hinge(80-86)-w-N^{AR}	495	579	772	70	175
AC-w-O^{24}	1723	1318	1059	1540	2482
Asp145-w-O^{24}	1402	1037	706	1115	2422
Glu51-w-O^{24}	0	8	5	1	0
Lys33-w-O^{24}	0	0	0	0	0
Gln131-w-O^{24}	825	317	57	631	32
Asn132-w-O^{24}					

[a] Water.
[b] N^{AR} pyridine nitrogen (Fig. 1).

by only the hinge segment. As expected, the hinge residues were the major contributors towards this kind of interaction due to their close proximity to N^{AR}. In case of CDK4-17 complex this water mediated interaction was quite strong in the first 2 and last 3 ns of the simulation with a residence time of 56% and 49%, respectively. Between 2 and 8 ns, this water mediated interaction resided for just 6% of the duration of the simulation. This evidently concedes the fact that during 2–8 ns of the simulation a direct interaction of $His95^{CDK4}$ with N^{AR} eliminates this water bridge interaction.

The inability of pyridine nitrogen to mediate significant water-mediated interaction when bound to CDK2 was very surprising. The significant difference in water-mediated interaction between CDK2-17 and CDK4-17 complex could be attributed to the difference in the orientation of pyridine aromatic ring with respect to the pyrido[2,3-*d*]pyrimidin-7-one core. The pseudo-dihedral angle between the aromatic rings of pyridine (A) and the core (B) were measured as a function of time (see Supplementary Material, Fig. S11), which indicates that pyridine aromatic ring, explores a range of conformation only in CDK2-17. In the initial stages of simulation the dihedral angle between the two aromatic rings A and B, in CDK2-17 remains at about −60° but then shoots towards the positive range indicating that it shifts more towards the hinge backbone. In the last 5 ns, it moves away from backbone and is oriented towards the aromatic ring of Phe82. The distance between N^{AR} and center of mass of the aromatic rings in Phe82 and His95 in CDK2-17 and CDK4-17, respectively, were also monitored as a function of time to identify the difference in the orientation of N^{AR} (Fig. 10). It indicates that in CDK2-17, the phenyl lies very close to the pyridine nitrogen while the corresponding His95 remains away from N^{AR} during 0–2 and 8–11 ns of the simulation when it does not directly interact with N^{AR}. Hence it can be argued that the lack of significant water mediated interaction in CDK2-17 is a direct outcome of the close proximity between N^{AR} and Phe82. In CDK4-17 complex, the His95 moves far apart from the pyridine during 0–2 ns and 8–11 ns of the simulation, enabling water molecules to approach the acceptor center N^{AR} to mediate bridged–water interaction. N^{AR} is closer to $Phe82^{CDK2}$ than ever before in the last 5 ns of the simulation, which also interestingly overlaps with the time when Lys89 intrudes the active site and severely hampers hydrogen bonding interaction with $Leu83^{CDK2}$. Taken together, it becomes clear that Lys89 causes lig17 to not only lose its ability to form hydrogen bond with carbonyl of $Leu83^{CDK2}$ but also prevents N^{AR} to hold bridge–water interaction with the protein. The inability of lig17 to club such interactions when bound to CDK2 would directly affect its affinity towards the protein.

In CDK2, Asp145-Lys33-Asp51 forms an essential catalytic triad (Asp158-Lys35-Glu56 in CDK4), which is an essential core for catalysis. Asp145 especially chelates the magnesium ion that coordinates with β and γ phosphates of ATP [20]. Hence, this region is at the center of the catalytic activity where γ phosphate transfer essentially occurs. So it is not surprising that several inhibitors of CDK2 target this site for extending its inhibitory activity [43]. For example, flavopiridol forms hydrogen bonding between its charged piperidine ring and both Lys33 and Asp145 of CDK2 [53]. Hymenialdisine also interacts with these residues by forming two hydrogen bonding interactions, one meditated by water, between guanidine ring and Asp145 [54]. Bound with this information we investigated whether this kind of interaction prevailed in the ligand under consideration. To begin with, we measured all possible water-mediated hydrogen bonding interaction of O^{24} with CDK4 and CDK2 and it was observed that these interactions were slightly better in the latter than former (74% against 67%). We observed a very strong tendency for the carbonyl oxygen (O^{24}) of pyrido[2,3-d]pyrimidin-7-one core to interact with $Asp145^{CDK2}$ via a water molecule. Lig17 interaction with $Asp145^{CDK2}$ through a water molecule was observed extensively in CDK2-17 than that with the corresponding residue, $Asp158^{CDK4}$

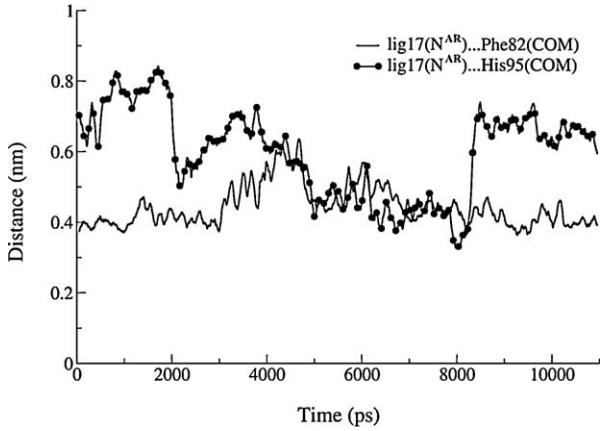

Fig. 10. The distance between N^{AR} and the center of mass (COM) of the aromatic ring, $Phe82^{CDK4-17}$ and $His95^{CDK4-17}$.

Table 3
Eigenvalues of the first 10 eigenvectors obtained from PCA.

Eigen-vector	CDK2		CDK217		CDK4		CDK417	
	EV[a]	CV[b]	EV	CV	EV	CV	EV	CV
1	9.44	49.56	8.67	43.45	6.11	27.62	9.23	40.76
2	1.60	57.93	2.66	56.80	2.76	40.10	3.25	55.14
3	1.05	63.44	0.99	61.76	2.29	50.46	1.61	62.28
4	0.55	66.33	0.79	65.74	1.59	57.65	1.02	66.80
5	0.47	68.78	0.61	68.79	0.87	61.57	0.57	69.33
6	0.37	70.73	0.44	71.01	0.77	65.04	0.52	71.61
7	0.36	72.59	0.40	73.02	0.68	68.12	0.43	73.51
8	0.30	74.18	0.29	74.47	0.52	70.48	0.41	75.32
9	0.29	75.72	0.26	75.77	0.43	72.41	0.31	76.69
10	0.24	76.97	0.22	76.88	0.33	73.91	0.28	77.93
Total	19.05		19.94		22.10		22.63	

[a] Eigenvalue of each component in the PCA.
[b] Cummulative variance of the eigenvalues.

Table 4
The distance between the two extreme projections of the segments considered (nm).

CDK2		Eigenvector 1		Eigenvector 2	
Segment	Residues	CDK2	CDK217	CDK2	CDK217
G-loop	10–18	0.31	0.15	0.23	0.26
Dis.-loop	37–44	0.08	0.40	0.43	0.32
α1	45–58	0.09	0.21	0.13	0.14
T1-loop	147–156	0.07	0.22	0.07	0.10
T2-loop	157–165	0.07	0.15	0.07	0.06
L14	220–247	0.08	0.15	0.11	0.15
Nter	1–80	0.11	0.09	0.08	0.07
Cter	89–298	0.04	0.03	0.03	0.03
CDK4		Eigenvector 1		Eigenvector 2	
Segment	Residues	CDK4	CDK417	CDK4	CDK417
G-loop	12–20	0.33	0.18	0.13	0.11
Dis.-loop	39–49	0.25	0.38	0.25	0.40
α1	50–67	0.14	0.10	0.14	0.07
T1-loop	160–169	0.17	0.08	0.12	0.09
T2-loop	170–177	0.13	0.10	0.12	0.10
L14	231–256	0.17	0.24	0.30	0.13
Nter	1–93	0.14	0.07	0.08	0.09
Cter	102–303	0.06	0.03	0.04	0.04

in CDK4-17. Asp145^{CDK2} was observed to mediate bridge water hydrogen bonding interaction for about 60% of the duration of simulation while Asp158^{CDK4} was able to maintain the same for a mere 16%. Asn145^{CDK4}-wat-O^{24} interaction seems to get better with simulation time and is able to mediate interaction for 16% of the simulation time. In our simulation of CDK2-17, Gln131 was observed to form a weak water-mediated interaction with ligand while on the other hand Glu144$^{CDK4-17}$ was the major contributor to mediate bridged water interaction with protein via O^{24}.

Protein functions are embedded in their structure and a major goal of protein simulation is to generate enough configurations of the system of interest to extract functionally relevant motions. One such tool for reducing the generated dimensionality of MD trajectory to an essential subspace encompassing few degrees of freedom, eliminating the positional fluctuations, is principal component analysis (PCA), also called essential dynamics [35]. Eigenvalues represent the amplitude of the correlated motions exhibited by the system while eigenvectors project the direction of these motions. In other words, each of the eigenvalues represents the percentage of total fluctuation demonstrated by the protein in the direction of the eigenvector during the MD simulation. Table 3 lists eigenvalues corresponding to the first ten eigenvectors extracted from the trajectories. The first ten eigenvectors are able to explain about 70% of the total relevant motions covered during the MD run. It is well documented that the first few eigenvectors capture the essential motions exhibited by the system during the MD run. In case of CDK2, the motions corresponding to the first eigenvector explain about 50% of the relevant motions spanned during the simulation while that in CDK2-lig17 is able to explain about 44%. In CDK4, the motion pertinent to the first eigenvector is able to explain still less, only 27% of the significant motions while CDK4-17 complex explains 41% of same during MD. Since the first eigenvector corresponds to the most significant motion exhibited by the system, the unliganded CDK4 seems to undergo less correlated motions than its complexed form while on the other hand unliganded CDK2 seems to exhibit exuberant motions than its complexed form.

Several trails of evidence support the high flexibility of G-loop and T-loop [45,46,55]. Since G-loop is at the entrance of the active site, evaluating its structural motion was perceived to be significant. In addition, the motion of the α1-helix as well as that of T-loop was also monitored. Analysis of PCA results highlighted T-loop region to get attracted towards N-terminal domain in some cases while the later part of the loop was found to move towards the C-terminal domain. Hence, we split the T-loop into two regions namely T1-loop and T2-loop to primarily highlight the difference in the two sections of T-loop. Disordered loop have earlier been investigated to be highly mobile [12] and the L14 loop was also considered since it shows ample fluctuations within the simulation time frame (Fig. 2). PCA results indicated G-loop to be the region exhibiting high fluctuations followed by the disordered and L14 loop (Table 4). The disordered loop and L14 loop lie at the extreme end of the N-terminal and C-terminal domains respectively and their motions were visualized to indicate the associative or dissociative motion of the two segments. The arrows in the figures (Fig. 11, S11 to S12) indicate the direction of motion of the segments, with the point of origin corresponding to the center of mass (COM) of the first frame. An arrow projecting the COM motion is drawn only when the segments move by a distance greater than 0.1 nm. The length of the arrow is proportional to the extent of movement of the particular segment.

The motion corresponding to the first eigenvector in case of CDK2 is concentrated on the G-loop that moves towards the C-terminal domain (Fig. 11). The N-terminal domain seems to follow the path of G-loop while the rest of the segments selected for evaluation seem to be unperturbed. The motions localized by the first eigenvector correspond well with loop regions of the protein. In CDK2-17 complex the G-loop motion is comparatively restricted but other segments begin to exhibit enhanced fluctuations. These motions correspond well with results previously reported for CDK2 in complex with roscovitine and olomoucine II [56]. They observed G-loop and flexible N-terminal loops move in a direction that leads to opening of the ATP-binding site. The T1-loop is observed to move away from the N-terminal while the second T-loop segment, T2-loop, shows an upward motion. Our finding suggests that in CDK2-17 complex, the primary motions of these loops converge into the active site of the protein resulting in a closure of the ATP pocket. The PCA results also indicate disordered and L14 loops to move in a direction facing each other. Also the α1-helix shows a feeble motion in a direction opposite to the catalytic center. The difference between the two CDK homologues is highlighted when the motion corresponding to the first eigenvector of CDK4 is compared to its CDK2 counterpart. The two segments, G-loop and disordered loop, indicated in the figure (Fig. 11), clearly highlights the N-terminal domain prefers to move away from the C-terminal domain in CDK4 and CDK4-17. In CDK4, the T1-loop and α1 helix moves towards the active site center while in CDK4-17, the T-loops does not exhibit significant motion. The disordered as well as L14 loops in both CDK4 and CDK4-17 migrates in opposite direction

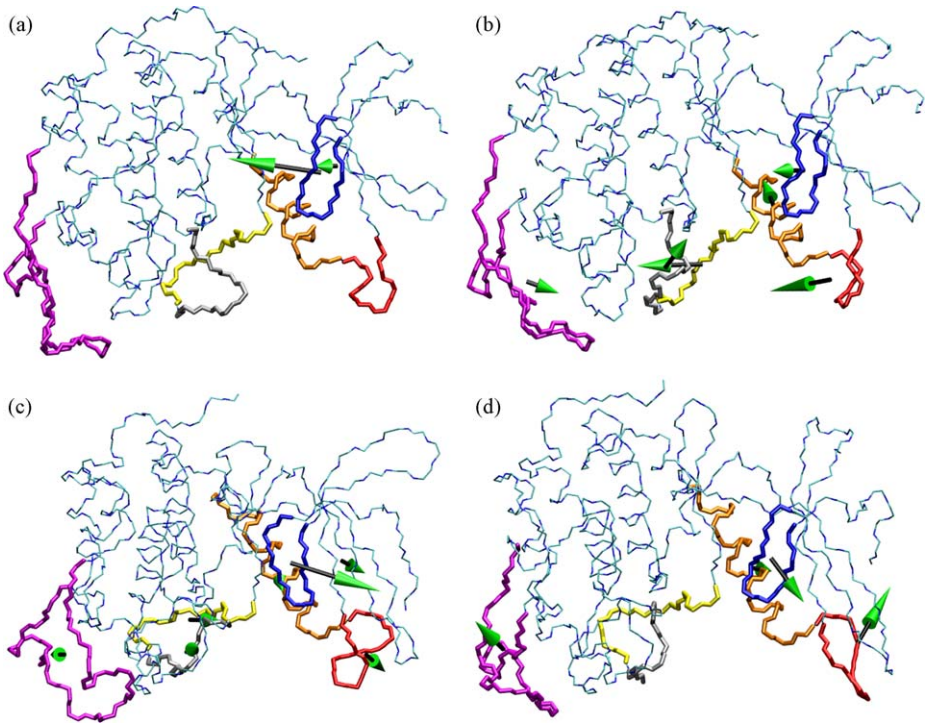

Fig. 11. Visualization of the dominant motions corresponding to first eigenvector in (a) CDK2, (b) CDK217, (c) CDK4, (d) CDK417. The color coded region corresponds to various segments of the protein: G-loop (blue), Disordered loop (red), α1-helix (orange), T1-loop (yellow), T2-loop (grey) and L14 (magenta). The backbone structure represents the initial conformation of the projection of the eigenvector. The arrow projects the direction of motion of the center of mass of the segments between the two extreme conformations explored by the protein corresponding to the eigenvector.

away from one another, quite opposite to what was observed in CDK2.

In the motion corresponding to the second eigenvector, the movement is concentrated on the disordered loop in CDK2 (see Supplementary Material, Fig. S12). In both CDK2 and CDK2-17, the disordered loop moves towards the G-loop while the L14 loop moves in a direction opposite to the ATP pocket. The G-loop in both the cases exhibits an upward motion probably aided by the inward movement of disordered loop. The T1-loop in CDK2-17 is found to move away from the N-terminal domain, similar to the motion corresponding to first eigenvector. The motion of the G-loop in CDK4 and CDK4-17 are in good agreement with those of CDK2 and CDK2-17, respectively. The big difference arises from the orientation of the movement of the disordered loop. In both CDK4 and CDK4-17 the disordered loop and L14 loop move in a direction opposite to each other while in CDK2 and CDK2-17 the disordered loop moves inwards towards the active site center.

4. Conclusions

A detailed perspective from an atomic viewpoint has been provided regarding the selectivity aspect of lig17 towards CDK4. The role of the positive charge on lig17 in determining its selectivity towards CDK4 over CDK2 has been rationally analyzed. The electronegative character of CDK4 active site in comparison to CDK2 has been found to exert a strong attraction towards the positively charged ligand. The presence of Lys89 in CDK2 has been found to create a significant stress for the positively charged lig17, directly affecting its conformation and consequently obstructing its hydrogen bonding abilities with the protein. Through a series of analysis it was possible to elucidate the mechanism by which the pyridine containing substituents provide interaction centers that enhances its affinity towards CDK4 than CDK2. The pyridine nitrogen was found to interact with the side chain of $His95^{CDK4}$ through hydrogen bonding interaction. It was also found to be capable of interacting with hinge segment of CDK4 through water mediated contacts. Both these observations have been elucidated to be crucial in determining its selectivity towards CDK4. In comparison to CDK2-17, G-loop in CDK4-17 exhibited enhanced mobility. Residue Tyr17 was particularly solvent exposed in both CDK4 and CDK4-17 while the corresponding Tyr15 in CDK2 and CDK2-17 was well buried. The distinct behavior of G-loop in CDK4-17 and the unique "flipping-out" mechanism observed in $Tyr17^{CDK4/CDK4-17}$ further broadens the diverse nature of two homologous CDKs. With the aid of PCA, dominant motions exhibited by the proteins have also been extracted. Most of the motions were concentrated on the N-terminal domain of the protein. The G-loop, disordered loop and L14 loop were observed to behave quite differently in CDK2 and CDK4. The functional and structural insights gained in the current study are expected to be very useful for structure guided drug designing.

Acknowledgements

N.M.M. is grateful to CSIR, New Delhi, India for the fellowship (SRF). N.G. is thankful to CSIR Mission Mode Program (CMM-0017) for the financial assistance. N.M.M. thanks tutors of GROMACS and NAMD workshop at WBUT and JNCASR respectively.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.01.008.

References

[1] K.J.T. Collins, N.P. Pavletich, The cell cycle and cancer, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 2776–2778.
[2] D. Morgan, Cyclin-dependent kinases: engines, clocks, and microprocessors, Annu. Rev. Cell Dev. Biol. 13 (1997) 261–291.

[3] M.M. Kasten, A. Giordano, pRb and the cdks in apoptosis and the cell cycle, Cell Death Differ. 5 (1998) 132–140.
[4] S. Ortega, M. Malumbres, M. Barbacid, Cyclin D-dependent kinases, INK4 inhibitors and cancer, Biochim. Biophys. Acta 1602 (2002) 73–87.
[5] S. Ortega, M. Malumbres, M. Barbacid, Cell cycle and cancer: the G1 restriction point and the G1/S transition, Curr. Genomic. 3 (2002) 245–263.
[6] M.W. Landis, B.S. Pawlyk, T. Li, P. Sicinski, P.W. Hinds, Cyclin D1-dependent kinase activity in murine development and mammary tumorigenesis, Cancer Cell 9 (2006) 13–22.
[7] Q. Yu, E. Sicinska, Y. Geng, M. Ahnstrom, A. Zagozdzon, Y. Kong, H. Gardner, H. Kiyokawa, L.N. Harris, O. Stal, P. Sicinski, Requirement for CDK4 kinase function in breast cancer, Cancer Cell 9 (2006) 23–32.
[8] P. Sridhar, N. Akula, N. Pattabiraman, Selectivity and potency of cyclin-dependent kinase inhibitors, AAPS J. 8 (2006) E204–E221.
[9] T. Honma, T. Yoshizumi, N. Hashimoto, K. Hayashi, N. Kawanishi, K. Fukasawa, T. Takaki, C. Ikeura, M. Ikuta, I. Suzuki-Takahashi, T. Hayama, S. Nishimura, H. Morishima, A novel approach for the development of selective CDK4 inhibitors: library design based on locations of CDK4 specific amino acid residues, J. Med. Chem. 44 (2001) 4628–4640.
[10] C. McInnes, S. Wang, S. Anderson, J. O'Boyle, W. Jackson, G. Kontopidis, C. Meades, M. Mezna, M. Thomas, G. Wood, D.P. Lane, P.M. Fischer, Structural determinants of cdk4 inhibition and design of selective ATP competitive inhibitors, Chem. Biol. 11 (2004) 525–534.
[11] K.A. Rossi, J.A. Markwalder, S.P. Seitz, C.H. Chang, S. Cox, M.D. Boisclair, L. Brizuela, S.L. Brenner, P.F. Stouten, Understanding and modulating cyclin-dependent kinase inhibitor specificity: molecular modeling and biochemical evaluation of pyrazolopyrimidinones as CDK2/cyclin A and CDK4/cyclin D1 inhibitors, J. Comput. Aided Mol. Des. 19 (2005) 111–122.
[12] H. Park, M.S. Yeom, S. Lee, Loop flexibility and solvent dynamics as determinants for the selective inhibition of cyclin-dependent kinase 4: comparative molecular dynamics simulation studies of CDK2 and CDK4, Chembiochem 5 (2004) 1662–1672.
[13] Y. Jiang, J. Zou, C. Gui, Study of a ligand complexed with CDK2/CDK4 by computer simulation, J. Mol. Model. 11 (2005) 509–515.
[14] N. Dasselew, P.V. Bharatam, 3D-QSAR and molecular docking study on bisarylmaleimide series as glycogen synthase kinase 3, cyclin dependent kinase 2 and cyclin dependent kinase 4 inhibitors: an insight into the criteria for selectivity, Eur. J. Med. Chem. 42 (2007) 1014–1027.
[15] D.J. Pratt, J. Bentley, P. Jewsbury, F.T. Boyle, J.A. Endicott, M.E. Noble, Dissecting the determinants of cyclin-dependent kinase 2 and cyclin-dependent kinase 4 inhibitor selectivity, J. Med. Chem. 49 (2006) 5470–5477.
[16] P.J. Day, A. Cleasby, I.J. Tickle, M. O'Reilly, J.E. Coyle, F.P. Holding, R.L. McMenamin, J. Yon, R. Chopra, C. Lengauer, H. Jhoti, Crystal structure of human CDK4 in complex with a D-type cyclin, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 4166–4170.
[17] T. Takaki, A. Echalier, N.R. Brown, T. Hunt, J.A. Endicott, M.E. Noble, The structure of CDK4/cyclin D3 has implications for models of CDK activation, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 4171–4176.
[18] P.L. Toogood, P.J. Harvey, J.T. Repine, D.J. Sheehan, S.N.V. Wel, H. Zhou, P.R. Keller, D.J. McNamara, D. Sherry, T. Zhu, J. Brodfuehrer, C. Choi, M.R. Barvian, D.W. Fry, Discovery of a potent and selective inhibitor of cyclin-dependent kinase 4/6, J. Med. Chem. 48 (2005) 2388–2406.
[19] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[20] P.D. Jeffrey, A.A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massague, N.P. Pavletich, Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex, Nature 376 (1995) 313–320.
[21] N.M. Mascarenhas, N. Ghoshal, Combined ligand and structure based approaches for narrowing on the essential physicochemical characteristics for CDK4 inhibition, J. Chem. Inf. Model. 48 (2008) 1325–1336.
[22] H. Lu, U. Schulze-Gahmen, Toward understanding the structural basis of cyclin-dependent kinase 6 specific inhibition, J. Med. Chem. 49 (2006) 3826–3831.
[23] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.
[24] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.
[25] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, R.P. Mee, Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes, J. Comput. Aided Mol. Des. 11 (1997) 425–445.
[26] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[27] A.W. Schuttelkopf, D.M. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 1355–1363.
[28] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordaon, J.H. Jensen, S. Koseki, K.A. Nguyen, S. Su, T.L. Windus, M. Dupius, J.A. Montgomery, General atomic and molecular electronic structure systems, J. Comput. Chem. 14 (1993) 1347–1363.
[29] Senda, N. http://winmostar.com/help_en.html.
[30] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comp. Chem. 18 (1997) 1463–1472.
[31] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. Dinola, J.R. Haak, Molecular-dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[32] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
[33] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[34] M. Gilson, K. Sharp, B. Honig, Calculating electrostatic interactions in bio-molecules: Method and error assessment, J. Comp. Chem. 9 (1988) 327–335.
[35] A. Amadi, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412–425.
[36] M. Otyepka, Z. Kríz, J. Koca, Dynamics and binding modes of free CDK2 and its two complexes with inhibitors studied by computer simulations, J. Biomol. Struct. Dyn. 20 (2002) 141–154.
[37] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[38] H.L. De Bondt, J. Rosenblatt, J. Jancarik, H.D. Jones, D.O. Morgan, S.H. Kim, Crystal structure of cyclin-dependent kinase 2, Nature 363 (1993) 595–602.
[39] G. Lolli, E.D. Lowe, N.R. Brown, L.N. Johnson, The crystal structure of human CDK7 and its protein recognition properties, Structure 12 (2004) 2067–2079.
[40] D.H. Brotherton, V. Dhanaraj, S. Wick, L. Brizuela, P.J. Domaille, E. Volyanik, X. Xu, E. Parisini, B.O. Smith, S.J. Archer, M. Serrano, S.L. Brenner, T.L. Blundell, E.D. Laue, Crystal structure of the complex of the cyclinD-dependent kinase CDK6 bound to the cell-cycle inhibitor p19INK4d, Nature 395 (1998) 244–250.
[41] U. Schulze-Gahmen, H.L. De Bondt, S.H. Kim, High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design, J. Med. Chem. 39 (1996) 4540–4546.
[42] E. Radzio-andzelm, J. Lew, S. Taylor, Bound to activate: conformational consequences of cyclin binding to CDK2, Curr. Biol. 3 (1995) 1135–1141.
[43] T.G. Davies, D.J. Pratt, J.A. Endicott, L.N. Johnson, M.E. Noble, Structure-based design of cyclin-dependent kinase inhibitors, Pharmacol. Ther. 93 (2002) 125–133.
[44] H. Lu, D.J. Chang, B. Baratte, L. Meijer, U. Schulze-Gahmen, Crystal structure of a human cyclin-dependent kinase 6 complex with a flavonol inhibitor, fisetin, J. Med. Chem. 48 (2005) 737–743.
[45] J.P.I. Welburn, J.A. Tucker, T. Johnson, L. Lindert, M. Morgan, A. Willis, M.E.M. Noble, J.A. d'Endicott, How tyrosine 15 phosphorylation inhibits the activity of cyclin-dependent kinase 2-cyclin A, J. Biol. Chem. 282 (2007) 3173–3181.
[46] I. Bartova, M. Otyepka, Z. Kriz, J. Koca, Activation and inhibition of cyclin-dependent kinase-2 by phosphorylation; a molecular dynamics study reveals the functional importance of the glycine-rich loop, Protein Sci. 13 (2004) 1449–1457.
[47] T.G. Davies, J. Bentley, C.E. Arris, F.T. Boyle, N.J. Curtin, J.A. Endicott, A.E. Gibson, B.T. Golding, R.J. Griffin, I.R. Hardcastle, P. Jewsbury, L.N. Johnson, V. Mesguiche, D.R. Newell, M.E. Noble, J.A. Tucker, L. Wang, H.J. Whitfield, Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor, Nat. Struct. Biol. 9 (2002) 745–749.
[48] S. Scheiner, T. Kar, J. Pattanayak, Comparison of various types of hydrogen bondsinvolving aromatic amino acids, J. Am. Chem. Soc. 124 (2002) 13257–13264.
[49] A. Jain, V. Ramanathan, R. Sankararamakrishnan, Lone pair . pi interactions between water oxygens and aromatic residues: quantum chemical studies based on high-resolution protein structures and model compounds, Protein Sci. 18 (2009) 595–605.
[50] B. Zhang, V.B. Tan, K.M. Lim, T.E. Tay, Significance of water molecules in the inhibition of cyclin-dependent kinase 2 and 5 complexes, J. Chem. Inf. Model. 47 (2007) 1877–1885.
[51] Z. Kríz, M. Otyepka, I. Bártová, J. Koca, Analysis of CDK2 active-site hydration: a method to design new inhibitors, Proteins 55 (2004) 258–274.
[52] M.R. Finlay, D.G. Acton, D.M. Andrews, A.J. Barker, M. Dennis, E. Fisher, M.A. Graham, C.P. Green, D.W. Heaton, G. Karoutchi, S.A. Loddick, R. Morgentin, A. Roberts, J.A. Tucker, H.M. Weir, Imidazole piperazines: SAR and development of a potent class of cyclin-dependent kinase inhibitors with a novel binding mode, Bioorg. Med. Chem. Lett. 18 (2008) 4442–4446.
[53] W.F. De Azevedo, H.J. Mueller-Dieckmann, U. Schulze-Gahmen, P.J. Worland, E. Sausville, S.H. Kim, Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2. a cell cycle kinase, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 2735–2740.
[54] L. Meijer, Chemical inhibitors of cyclin-dependent kinases, Trends Cell Biol. 6 (1996) 393–397.
[55] A.A. Russo, P.D. Jeffrey, N.P. Pavletich, Structural basis of cyclin-dependent kinase activation by phosphorylation, Nat. Struct. Biol. 3 (1996) 696–700.
[56] I. Bártová, J. Koca, M. Otyepka, Functional flexibility of human cyclin-dependent kinase-2 and its evolutionary conservation, Protein Sci. 17 (2008) 22–33.