Analysis of CDK2 Active-Site Hydration: A Method to Design New Inhibitors

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · MAY 2004

Impact Factor: 2.63 · DOI: 10.1002/prot.20026 · Source: PubMed

4 AUTHORS, INCLUDING:

Zdenek Kriz
Masaryk University
17 PUBLICATIONS **312** CITATIONS
SEE PROFILE

Michal Otyepka
Palacký University of Olomouc
180 PUBLICATIONS **4,344** CITATIONS
SEE PROFILE

PROTEINS: Structure, Function, and Bioinformatics 55:258–274 (2004)

Analysis of CDK2 Active-Site Hydration: A Method to Design New Inhibitors

Zdeněk Kříž,[1] Michal Otyepka,[2] Iveta Bártová,[1] and Jaroslav Koča[1]*
[1]*National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic*
[2]*Department of Physical Chemistry, Faculty of Science, Palacky University, Olomouc, Czech Republic*

ABSTRACT **The interactions between the protein and the solvent were analyzed, and protein regions with a high density of water molecules, as well as structural water molecules, were determined by using molecular dynamics (MD) simulations. A number of water molecules that were in contact with the protein for the whole trajectory were determined. Their interaction energies and hydrogen bonds with protein residues were analyzed. Altogether, 39, 27, 49, and 32 water molecules bound to the protein were found for trajectories of the free CDK2, CDK2/ATP, CDK2/roscovitine, and CDK2/isopentenyladenine complexes, respectively. Positions of observed water molecules were compared with X-ray crystallography data. Special attention was paid to water molecules in the active site of the enzyme, and especially to the deep pocket, where the N9 roscovitine side-chain is buried. Exchange of active-site water molecules with bulk water through the tunnel from the pocket was observed. In the CDK2/isopentenyladenine complex simulation, two water molecules that arrange interaction between the inhibitor and the enzyme via an H-bond were observed. Two stable water molecules in the trajectory of the free CDK2 were found that occupy the same position as the nitrogens N3 and N9 of the isopentenyladenine or N1 and N6 nitrogens of the adenosine triphosphate (ATP). The positions of structural water molecules were compared with the positions of substrate polar groups and crystallographic water molecules found in the Brookhaven Protein Data Bank for various CDK2 complexes. It was concluded that tracing tightly bound water molecules may substantially help in designing new inhibitors. Proteins 2004;55:258–274.**

Key words: cyclin-dependent kinase; ATP; roscovitine; isopentenyladenine; molecular dynamics; hydration of proteins; structural water molecules

INTRODUCTION

Water molecules play an essential role in the structure and dynamics of biological systems. In particular, the solvation of proteins is involved in most biological processes.[1] They act as a bridge between secondary structural elements.[2] Water is known to contribute significantly to the stability of biomacromolecules and to play a crucial role in molecular association. Buried water molecules are involved in local structural stabilization of proteins and are therefore termed "structural water."[3] They are usually characterized by long residence time and very restricted movement.[4] Water molecules in the binding interface can provide useful information for drug design. Strategies for the design of new inhibitors, which take into consideration structural water molecules, include either saturating the water position with appropriate hydrogen-bond donors and acceptors to achieve maximum complementarity or replacing it to gain in affinity through entropic effect.[5]

X-ray crystallography, NMR spectroscopy, and neutron diffraction are typical experimental methods to analyze water molecules at the atomic level. However, high-resolution structures are usually preferred for a reliable analysis of structural aspects of water associated with a host protein.[6] In the absence of high-resolution experimental data, theoretical studies, such as molecular dynamics (MD) simulations, represent complementary methods of locating water positions and understanding the dynamics and energetics of these water molecules.[7–10]

The cyclin-dependent kinases (CDKs) play a significant role in eucaryotic cell cycle regulation. Their sequential activation ensures the correct timing and ordering of events required for cell cycle progression. The CDKs activate host proteins through phosphorylation on serine or threonine using adenosine triphosphate (ATP) as a phosphate donor. The activity of CDKs is extensively regulated by association with regulatory subunits (cyclins) and by specific phosphorylation at a positive site (Thr160 in CDK2) or negative sites (Thr14 and/or Tyr15 in CDK2), or by association with native inhibitors.[11] The CDK2 associates with cyclin E to promote progression through the G1 phase and subsequently relocates into a complex with cyclin A that is essential for entry into the S phase. Selective CDK inhibitors would be invaluable tools for structural, cell, and molecular biologists working in the cell cycle field. Furthermore, the characterization and

Grant sponsor: Grant Agency of the Czech Republic; Grant number: 201/98/K041.

*Correspondence to: Jaroslav Koča, National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic. E-mail: jkoca@chemi.muni.cz

Received 9 June 2003; Accepted 14 September 2003

Published online 27 February 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20026

development of low-molecular-weight CDK inhibitors, which could restore normal cell cycle control in tumor cells, would also have considerable potential in cancer therapeutics.[12]

To date, only structures of CDK2, CDK5, and CDK6 solved by X-ray crystallography have been reported. The X-ray structures of the CDK2 complexed with ATP,[13] purvalanol,[14] staurosporine,[15] olomoucine and isopentenyladenine,[16] roscovitine,[17] flavopiridol,[18] quinazoline,[19] indirubin,[20] and many others have been published. Comparison of the X-ray structures for CDK2/cyclin A/ATP,[16] CDK2/cyclin A/p27^{Kip1} (native protein inhibitor),[21] CDK2/cyclin A and T160pCDK2/cyclin A complexes,[22] and CDK2 alone[13] reveals only limited differences in the vicinity of the ATP adenine binding site between different activation states of CDK2. This suggests that information derived from monomeric CDK2/inhibitor structures can help us to design inhibitors against the active binary complex.

All known eucaryotic protein kinases exhibit a conserved catalytic core domain with an adenosine 5′-triphosphate (ATP) binding site, which is often targeted in drug discovery programs. It is well known how the ATP binding site of CDK2 can accommodate structurally diverse inhibitor types. However, as ATP is common to kinases and other proteins, inhibitors of CDK2 competitive with ATP are also active in other proteins, which is reflected by the toxicity of these drugs. For the design of a selective inhibitor with high activity, it is necessary to describe all specific interactions of the ligand in the active site including those with solvent molecules.[23] Information obtained from MD simulations, especially calculated interaction energies and positions of structural water molecules in the active site, helps us to understand the specificities of the protein–ligand binding.

The aim of this work is to present analyses of structural water molecules determined from MD studies performed on free CDK2 and its three complexes: with native substrate ATP, with the nonselective inhibitor isopentenyladenine, and with the highly selective inhibitor roscovitine. MD can supplement information from X-ray crystallography mainly concerning the dynamics of the system and its behavior in solution. Calculated interaction energies based on MD results are more realistic than energies based on calculations from the X-ray structure due to relaxation of the system in its native environment and the averaging of energies over many single frames. In this article, we attempt to provide new insight into the CDK2 active site and to broaden the scope for rational drug design.

METHODS

All MD simulations presented in this work were carried out using the AMBER-6.0 program package[24] and the Cornell et al.[25] force field with the TIP3P water model. The X-ray structures were used as starting data for MD simulations. The free CDK2 and CDK2/ATP crystal structures were obtained from the Brookhaven Protein Data Bank (PDB codes: 1hcl and 1hck). The X-ray structures of the cdk2/roscovitine and cdk2/isopentenyladenine complexes were kindly provided by Prof. S.-H. Kim (University of California, Berkeley CA). The resolutions of X-ray data were 1.8 Å (free CDK2), 1.9 Å (CDK2/ATP), 2.4 Å (CDK2/roscovitine), and 1.8 Å (CDK2/isopentenyladenine), respectively.

Molecular Dynamics Simulation

First, the hydrogen atoms of unionizable residues were added to the X-ray structures using the PROTONATE program.[24] The protonation states of ionizable residues were set, and hydrogen atoms were added by the WHATIF[26] program suite, together with full optimization of the hydrogen-bond network. The used method is implemented in HB2NET module.[27] The method is based on special force field (H-bond force field). It places the hydrogen atom onto ionizable residues and optimizes its position to create a maximum number of H-bonds between residues. It is also able to "flip" the terminal dihedral angle of asparagine and glutamine to optimize the possibility of H-bond. The side-chain of histidine is also able to "flip" within this force field. However, in all the above changes, a penalty is put on energy value. In our case, all basic residues on the surface that may interact with water molecules were protonated; also, all acidic residues on the surface were deprotonated. This should well mimic the situation at physiological pH; therefore, it is likely that our results are realistic. Also, all histidines on the surface of the protein were fully protonated; considering internal histidines, only His125 was set as δ-protonated. Crystallographic water molecules were kept in the structure.

The all-atom model was neutralized by adding 13 chloride counterions and was immersed in a rectangular TIP3P water box. This was carried out using the EDIT program of the AMBER package.[24] The initial box size was $82 \times 75 \times 72$ Å for free CDK2 simulation and nearly the same for the simulations of CDK2 complexes. The system was at initially relaxed and slowly heated before trajectory production. First, the position of bulk water molecules and counterions was optimized using minimization procedure of SANDER program followed by 3-ps MD simulation on these residues. Then the optimization of hydrogen atoms of the protein was performed. The next procedure was to relax the position of protein side-chain and backbone using series of minimizations with decreased restrains applied on the backbone atoms. The last procedure was the 50 ps length MD simulation, with slow heating of the system from 10–298 K. The unrestrained MD was then run on the system. The molecular mechanics (MM) energy, density, and volume of the system and root-mean-square deviation (RMSD) of the protein to X-ray structure were monitored. The production phase was started after the density of the system became balanced around 1.0 g/cm^{-3}, temperature balanced around 298 K, and calculated RMSD was also stable. The size of the simulation box was $81 \times 72 \times 69$ Å for free CDK2 simulation at this time. Detailed description of the trajectory preparation can be found in one of our recent articles.[28] After the equilibration phase, a 1-ns trajectory was produced for each system. The coordinates were saved each picosecond. All energy minimization and

TABLE I. Statistical Data Concerning Water Molecules in the Molecular Dynamics Simulations

Trajectory	Number of crystallographic water molecules	Total number of water molecules in simulations	Number of tightly bound water molecules	
			Total	In the active site[a]
Free CDK2	180	12075	39	16 (23)
CDK2/ATP	108	11743	27	12 (10)
CDK2/ROSC	82	12525	49	15 (9)
CDK2/IPA	99	12020	32	11 (8)

[a]Numbers in parentheses refer to the number of water molecules in the CDK2 active site, detected by X-ray crystallography.

MD simulations were carried out using the SANDER program of the AMBER package.

The MD simulations were performed using the following conditions. A 2-fs timestep was used in all simulations, and the particle-mesh Ewald (PME)[29] method to treat electrostatic interactions was employed. The simulations were run under the periodic boundary conditions in the NPT ensemble at 298 K with Berendsen temperature coupling and constant pressure (1 atm) with isotropic molecule-based scaling.[30] The SHAKE algorithm,[31] with a tolerance of 10^{-5}, was applied to fix all bonds containing a hydrogen atom, and the nonbond pair list was updated every 10 steps. A 9.0-Å cutoff was applied.

Calculation of Solvent Density Distribution

All rotational and translational movements of the simulation box were removed before analyses. The water molecule coordinates were transformed, taking the periodic boundaries into account. The coordinates of water oxygen atoms were mapped onto the three-dimensional (3D) rectangular grid with a 0.5-Å grid step, producing an average 3D density distribution. This was carried out using the PTRAJ program.[24] The particular choice of the grid step is a compromise between the uncertainty in location of the density features and statistical error in the local density value that arises due to a lower number of counts in each grid cell. This methodology was developed in M. Pettitt's group (and is described, for example, in their article[32]). The water density map was fitted onto the average structure, and regions with a density larger than 140% of the bulk water density were localized. This was performed using the MIDAS PLUS program.[33]

Localization of Structural Water Molecules

Here, we were interested in water molecules that have longer contact with the protein (called "structural water molecules"). By definition,[4] these water molecules are strictly bonded to protein and show small fluctuations. One descriptor of fluctuations is mean-square displacement (MSD) value, which, being smaller than 1Å,^{2} is a characteristic for structural water molecules. Since the PTRAJ and MIDAS procedures also localize regions with high fluctuation of water molecules, we had to employ the new programs RESTIME and SURFTIME[34] for the detection of structural water molecules. Moreover, our inspection is somewhat extended, as it is based on an analysis of van der Waals contacts between water molecules and protein atoms. It means also that water molecules with much larger MSD will result from our analysis. We also show in this article that these water molecules may serve to locally stabilize proteins and, therefore, may be called structural as well.

A key characteristics of structural water molecules is the interaction time T_i (w,j) of the water molecule w with the protein atom j. In time t, it is calculated using Eq. (1):

$$T_i(w,j) = step \cdot \sum_{0}^{Nsteps} P(w,j)_t. \quad (1)$$

Here, the binary function, $P(w,j)_t$ is assigned by the value of 1 if the distance between the water molecule w oxygen atom and the protein atom j is smaller than the sum of Van der Waals radii of the oxygen atom and the atom j increased by a parameter RFACTOR. Otherwise, $P(w,j)_t$ is set to zero. The parameter *step* is a period by which snapshots of the trajectory are saved (1 *ps* in our case), and *Nstep* is the number of snapshots.

The RFACTOR was set to 2 Å, which allows for some limited fluctuations of water molecules around protein atoms. We considered as structural water molecules with interaction time larger than 950 ps within 1000 ps long trajectories. The value was selected because there was a gap for lower values up to 800 ps. We have also compared this value with the residence times of water molecules in the first hydration shell of ions. We have found that this is true in the case of monovalent ions in the range from 10 to 30 ps at 300 K.[35,36] In our case, we take the residence times of structural water molecules larger than 950 ps, which is in general agreement with experimental data for proteins.[37] The timescale at which water exchanges with other solvent molecules has been observed in the range of 10–50 ps for water bound in protein surface crevices, to 0.1–1.0 ns for strongly bound waters, and nanoseconds to milliseconds for interior waters.[4,38,39]

These water molecules were then analyzed using standard AMBER procedures. The H-bond analyses were carried out using the CARNAL program of the AMBER-6.0 package.[24] The structural water molecules in the active site of the enzyme were then analyzed from the interaction energy point of view. The energy decomposition between these molecules and amino acids of the active site was calculated using the ANAL program of the AMBER-6.0 package.[24]

All MD simulations were carried out on two 700 MHz processors of the SGI 1200 PC cluster. All analyses were

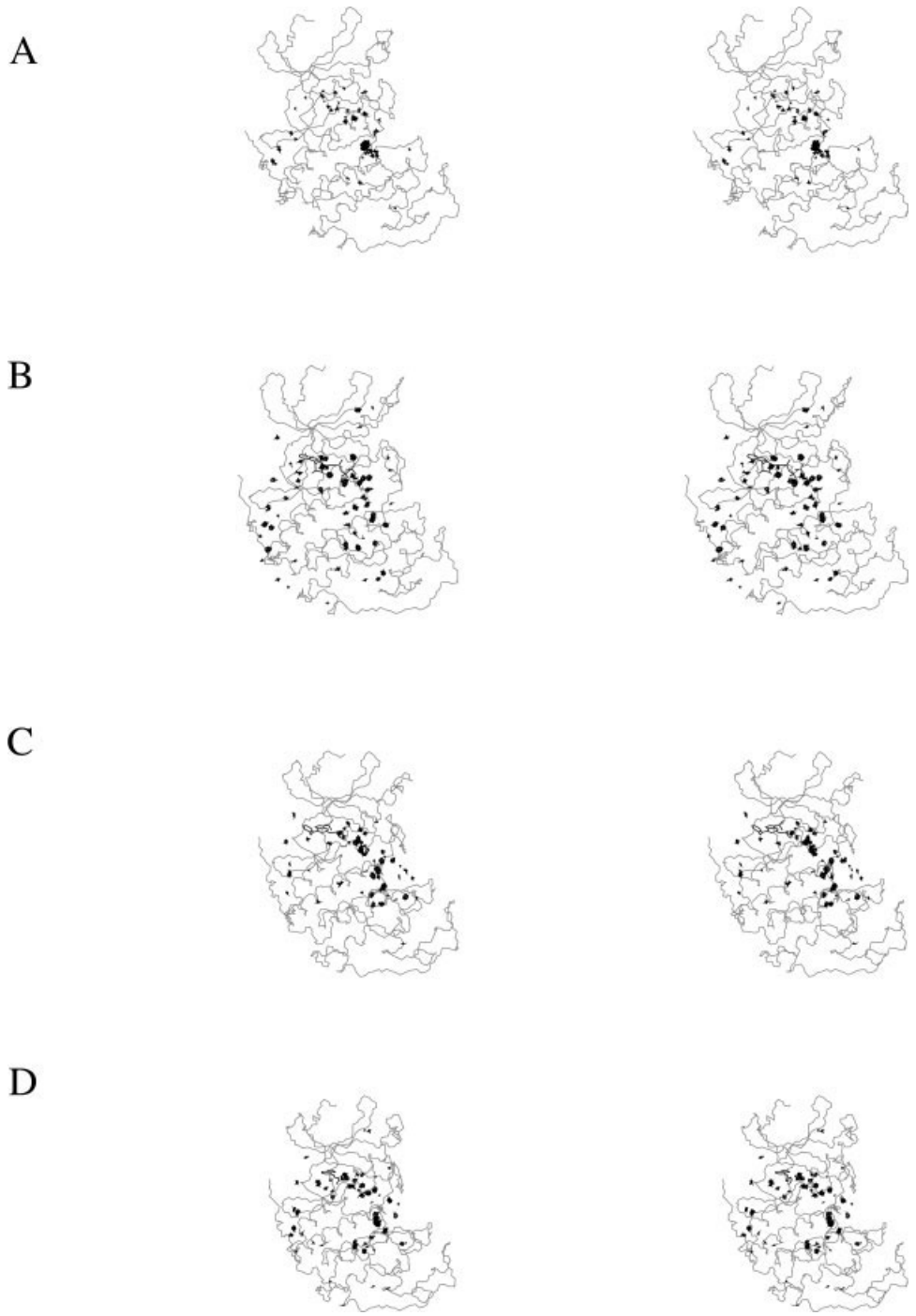


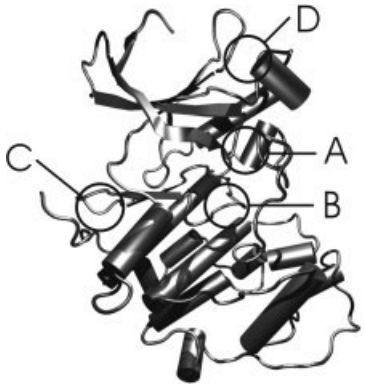
Fig. 1. Stereo view of regions with more than 140% bulk water density in the CDK2: (**A**) free CDK2, (**B**) CDK2/ATP, (**C**) CDK2/roscovitine, (**D**) CDK2/isopentenyladenine complexes.

performed on SGI Indigo workstations and PCs running Unix operating systems.

RESULTS AND DISCUSSION

We discussed the stability of produced trajectories and conformational behavior in a previous article.[28] The basic data about the number of water molecules are summarized in Table I.

The data in Table I shows that the total number of water molecules included in each simulation is comparable. The number of water molecules detected by X-ray crystallography depends on the structure hydration model used. As seen from Table I, the number of structural water molecules in the active site of the enzyme is comparable for simulations of the CDK2 complexes. For the free CDK2 simulation, the cavity is



Scheme 1. Scheme of the CDK2 secondary structure with highlighted solvated regions.

more open than in the crystal and will therefore be filled by more structural water molecules.

Hydrated Regions of the CDK2

The maps of regions with a high density of water molecules are depicted in Figure 1 (for notation, see also Scheme 1). The maps demonstrate that the hydrated regions are very similar for all simulations, and the presence of the ligand in the active site does not significantly affect the behavior of solvent around the enzyme. Figure 1 shows the three most hydrated regions. The first region (A) is in the CDK2 active site (around the αL12 helix, residues 148–151). The second region (B) surrounds the loop between the 3/10-1, α3 (residues 122–130), and the 3/10-2 helices (residues 182–198). The third region (C) is around the loop between the β6 and β7 strands (residues 135–141). As expected, the polar phosphate group of the ATP and the polar side-chain at the C2 position of the roscovitine inhibitor increase remarkably the density of water molecules in the active site. The differences between maps in the case of free CDK2 and CDK2/X complexes in the active site part could be explained by presence of polar atoms of ligands and by reorientation of amino acid side-chains in the active site. The ligand in the active site affects loops near the active site (G-loop and T-loop). It causes a partial opening of the active site so more water molecules can enter the cavity. Additionally, reorientation of the side-chains in regions C and B (Scheme 1) of the free CDK2 and CDK2/ligand complexes can also cause different solvation.

Structural Water Molecules

The localization of structural water molecules found in all MD simulations is depicted in Figure 2, and comparison of their positions with positions of X-ray-detected water molecules is shown in Figure 3.

Large clusters of structural water molecules mirror the previous results obtained from the maps of hydrated regions, especially in the case of the area surrounding the loop between the 3_{10}-1, α3, and 3/10-2 helices (region C). A large cluster of structural waters was found, as expected, in ribose and phosphate pockets of the active site. The detailed analysis also uncovers other regions with just a few structural water molecules. One of them is located between the β4, β5 strands and the PSTAIRE motif (region D).

There are many regions with only one structural water molecule, which is usually important for the structural stability. Such water molecules cannot be found by procedures based on tracing regions with a high density of water. There are also some areas (see Fig. 1) that are occupied by one or two structural water molecules frequently exchanged within a network of surrounding water molecules.

The structural water molecules detected by MD simulation are in similar positions to the X-ray detected water molecules. The best correspondence was found for free CDK2. The water molecules found on the surface of the protein by X-ray crystallography are relatively unstable, and bulk water molecules frequently exchange them. The number of X-ray detected water molecules is quite different in the case of CDK2/roscovitine and CDK2/isopentenyladenine complexes in comparison with free CDK2 and the CDK2/ATP complex. This discrepancy could be explained either by using different refinement procedures in the X-ray data analysis or by possible differences between the X-ray and solvent-relaxed structures.

A detailed inspection of Figure 3 shows that the region between the α1 helix and the loop between the β4 and β5 strands (region D) is hydrated by 3 water molecules in the case of free CDK2, 2 water molecules in the complex CDK2/ATP, and 4 water molecules and 2 water molecules in the cases of complexes CDK2/roscovitine and CDK2/isopentenyladenine, respectively. X-ray crystallography detects only one or no water molecule (CDK2/isopentenyladenine complex) here. The reason could be that the X-ray structures show high temperature factors for atoms in this region, and in some structures, the region was even excluded from the refinement due to the low electronic density found by X-ray.

Structural Water Molecules in the CDK2 Active Site

Free CDK2

Water molecules location. A detailed picture of the location of structural water molecules in the active site is shown in Figure 4. A comparison of the location of X-ray and MD structural water molecules is shown in Figure 5.

The analysis of the MD simulation data revealed 16 structural water molecules in the free CDK2 active site. Six of them are concentrated in the ribose pocket of the active site (waters 319, 320, 363, 377, 394, 6795). Only one of them (6795) is a bulk water molecule. Four water molecules (316, 317, 323, 418) create the cluster between the Lys33 side-chain and the glycine-rich loop (G-loop; residues 11–16). These water molecules were detected by X-ray crystallography, but only two of them kept their original position (316, 418). The position of the remaining ones (317, 323) changed during the simulation to create a stable cluster with stronger H-bonds to each other and to amino acids in the active site.

The next water cluster (348, 403, 412, 427) in the free

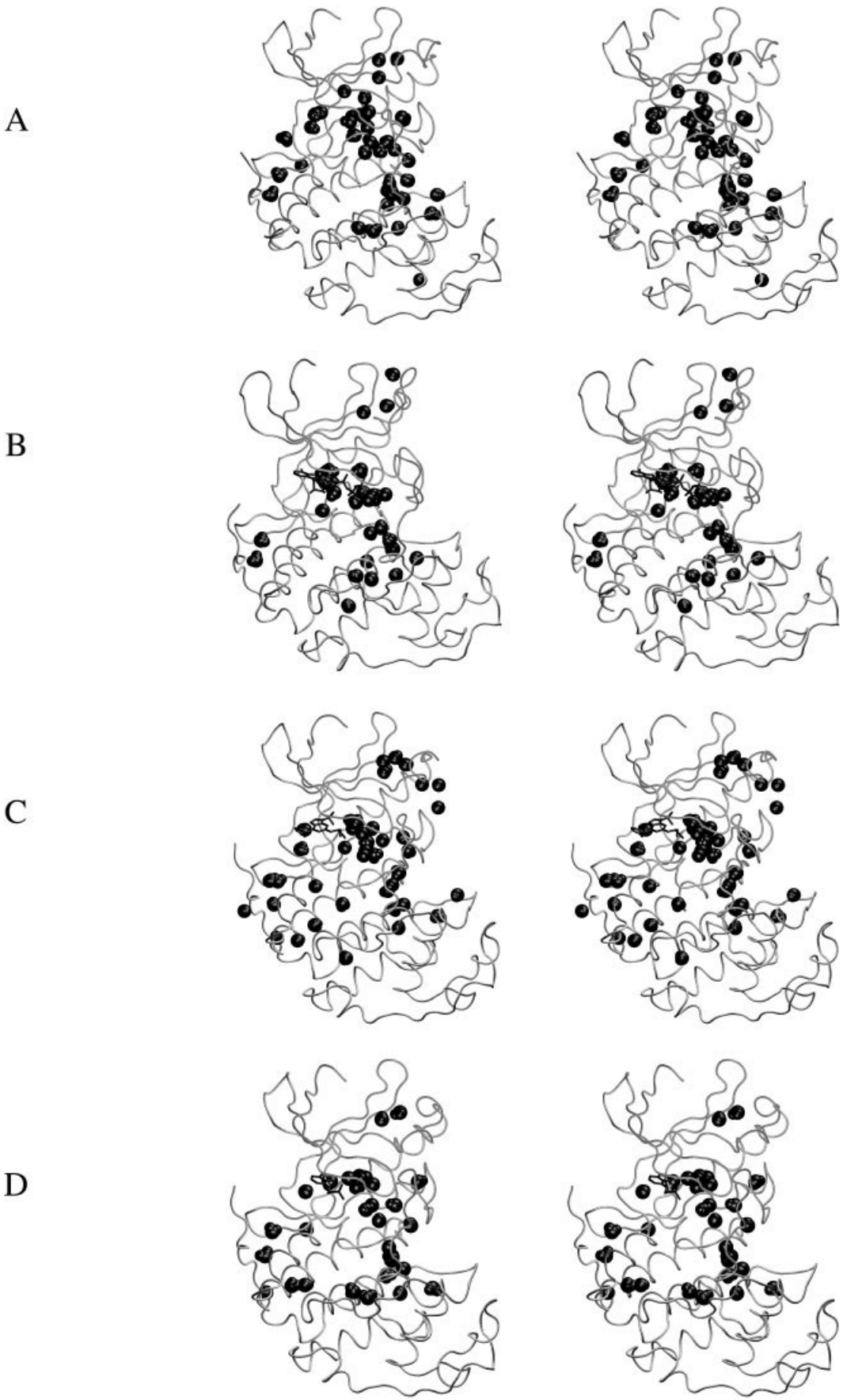


Fig. 2. The location of structural water molecules found by MD simulations: (**A**) free CDK2, (**B**) CDK2/ATP, (**C**) CDK2/roscovitine, (**D**) CDK2/isopentenyladenine complexes.

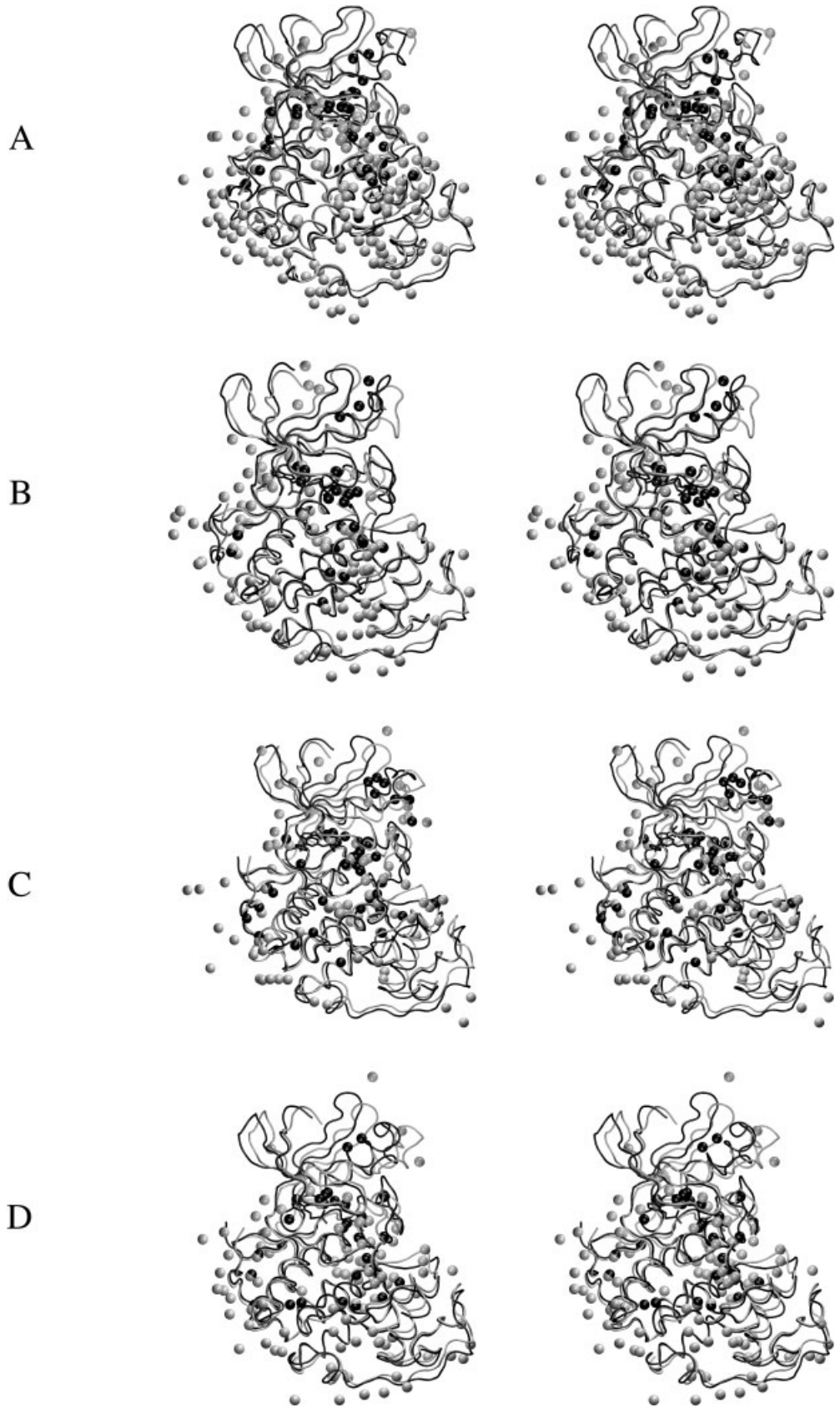


Fig. 3. The comparison of structural water molecule positions in the X-ray structure (gray) and in the average structures from equilibrated parts of MD trajectories (black). Water molecules are represented by the position of oxygen atoms. (**A**) free CDK2, (**B**) CDK2/ATP, (**C**) CDK2/roscovitine, (**D**) CDK2/isopentenyladenine complexes.

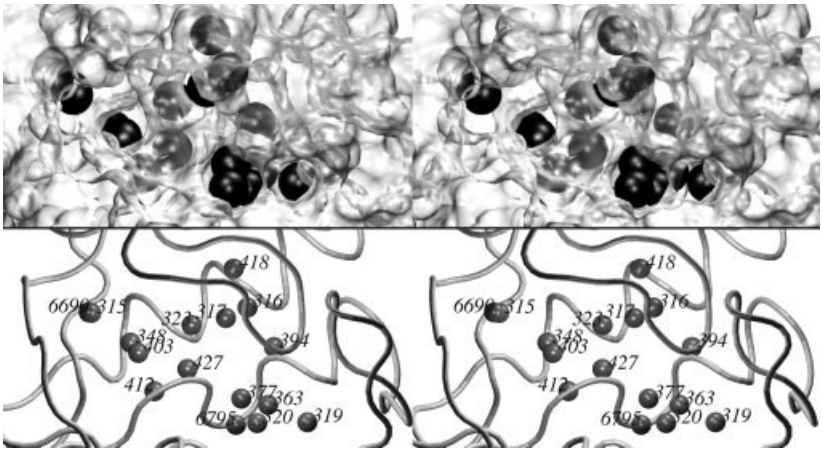


Fig. 4. The location of X-ray and MD structural water molecules in the CDK2 active site. The active site of CDK2 is modeled as a solvent-accessible surface; water molecules are in the CPK model (**top**). For better clarity, the same situation is pictured without the surface, and the positions of water molecules are represented by oxygen atoms (balls; **bottom**). Water molecules numbered as 480 or lower are those detected also by X-ray crystallography.

CDK2 active site is positioned in a shallow cavity created by Leu54, Lys56, Ile63, Val64, Ala144, Asp145, and Phe146. It is interesting that this cluster is located in proximity of hydrophobic residues. This is because of its stabilization by interactions with the polar backbone of the protein. The above-mentioned water molecules were also detected by X-ray crystallography, and their positions are relatively stable for the whole trajectory.

The last well-solvated region of the active site lies in the purine pocket (the hinge region created by residues 81–84) and is filled by 2 water molecules (315, 6690). This area is solvated only in free CDK2, because the nitrogens of the purine rings replace the positions of these water molecules in CDK2/purine inhibitor complexes. These water molecules create hydrogen bonds with the NH group of the Leu83 (6690) and carbonyl oxygen of the the Glu81 backbone (315). The positions of these 2 water molecules are the same as the positions of the N6 and N1 ATP nitrogens or N3 and N9 isopentenyladenine nitrogens in corresponding CDK2 complexes (Fig. 6). The positions of these 2 water molecules are also very close to the positions of the N2′, N3, and N9 atoms of roscovitine. The observation that structural water molecules in the free CDK2 are replaced by polar groups of ligand molecules in the complexed protein implies that the information about the location of structural water molecules is useful for rational drug design.

A detailed analysis of the X-ray structure shows 3 water molecules in this region. Only one of them (315) was detected as structural in the MD simulation. The second one was replaced by a bulk water molecule (6690) in the relaxation phase of the simulation, and it remained stable for the whole trajectory. The position of the third X-ray water molecule was not detected as a position of a structural water molecule.

Interaction energies. The calculated interaction energies between stable water molecules and CDK2 active site residues, together with the results of the H-bond analysis, are summarized in Table II. Some water molecules (315, 323, 403, 427, and 6690) interact with only one CDK2 residue during the whole trajectory. Several water molecules interact with two residues (316, 317, 377, 394, 412, and 6795). The remaining water molecules move slightly along neighboring residues and create alternate hydrogen bonds with a number of residues. Representatives of this category are waters 319, 320, 348, 363, and 418.

The hinge region (residues 81–84) is solvated by water molecules of the first category (interacting with just one atom). The interaction energies of water molecules in this region of the active site are −7.1 kcal/mol (water 315 and Glu81) and −4.8 kcal/mol (water 6690 and Leu83). The ribose pocket is filled by 4 water molecules: 316, 317, 323, and 418. The detailed analysis shows that the water 316 creates a bridge between the backbone of the Tyr15 and Gly16, and contributes to stability of the enzyme structure. Water molecule 317 interconnects the Lys33 and Asp145 side-chains. The high interaction energies calculated for this water molecule are caused by interactions with the charged side-chains of lysine and aspartic acid.

The phosphate pocket of the active site is solvated by a cluster of 4 water molecules (320, 363, 377, and 6795). A similar situation as for water 317 was observed for water 320, which stabilizes the orientation of the Asp127 and His125 side-chains. Its weak interactions with other CDK2 residues was also observed. The water molecule 363 moves along residues His125, Asp145, Gly127, and Ala149 during the simulation.

Detailed analysis of the interactions within the hydrophobic pocket of the active site (348, 403, 412, 427) shows that hydrogen bonds are created between these water molecules and the backbone atoms of the amino acids.

CDK2/ATP complex

Water molecules location. The presence of the substrate is reflected by the lower number of structural water molecules in the active site. The arrangement of these

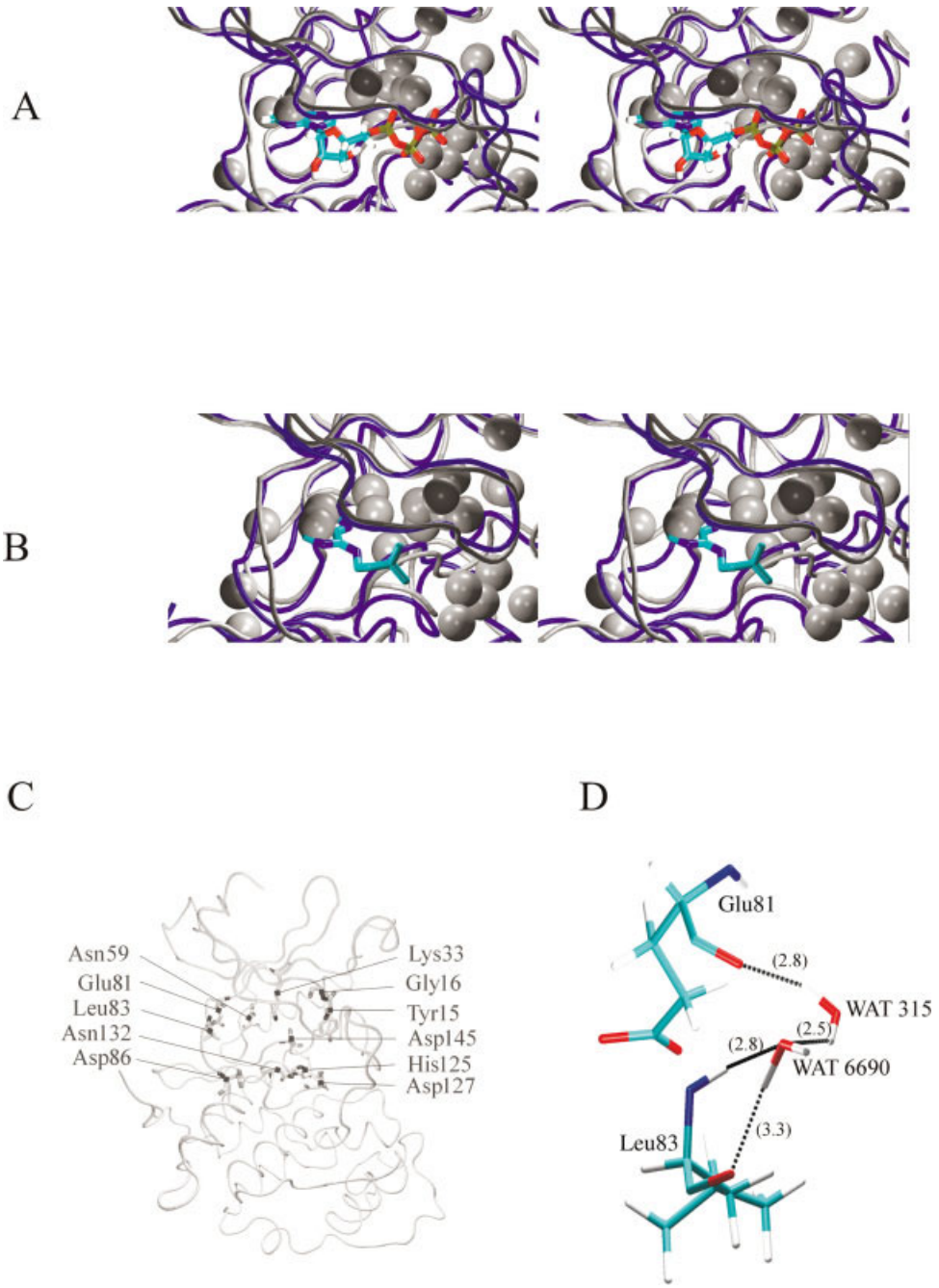


Fig. 6. Superimposition of CDK2/ATP complex (**A**) and CDK2/isopentenyladenine complex (**B**) with structural water molecules found for free CDK2. It is seen that structural water molecules in the free CDK2 occupy the same positions as the oxygen or nitrogen atoms of the ligands in the protein. (**C**) The position of the key residues of the CDK2 active site. (**D**) The detailed view of two structural water molecules (315 and 6690) bonded to Glu81 and Leu83 within the free CDK2 active site. Numbers in parenthesis show mean H-bond distances.

water molecules is shown in Figure 7. The ATP ribose is solvated by 5 water molecules (314, 315, 340, 379, and 381), while another water molecule (313) interacts with the G-loop. The ATP phosphate part is solvated by 5 water molecules (312, 330, 391, 396, and 7327). Comparison of the location of X-ray and MD structural water molecules is shown in Figure 8. Only 2 bulk water molecules became structural water molecules in the active site. One of them (3987) exchanged position with X-ray water and interacted with the ASP 86 side-chain and with ATP-ribose oxygen. The second one (7327) interacted with the ATP beta phosphate. It was also observed that the water cluster around the ATP gamma phosphate was restructured.

Interaction energies. The interaction energies and results of the H-bond analyses are summarized in Table III. Two water molecules, which interact with only one residue of the protein, were found. Structural water 314 solvates the Gln131 backbone, and water 315 interacts

TABLE II. Interaction Energies and H-Bond Analyses Calculated for MD Structural Water Molecules in the Free CDK2 Trajectory

Residue	Occupancy [%]	Distance [Å]	H-bond angle [degrees]	Interaction energy [kcal/mol]		
				Electrostatic	VdW	Total
			Water 315			
Glu81 (O)	99.2	2.8 ± 0.2	12.9 ± 7.4	−8.1 ± 1.7	1.0 ± 1.1	−7.1 ± 1.1
			Water 316			
Tyr15 (N-H)	99.4	2.9 ± 0.2	18.6 ± 7.9	−4.3 ± 1.3	0.3 ± 0.9	−3.9 ± 1.0
Gly16 (N-H)	83.8	3.3 ± 0.3	37.4 ± 11.5	−0.4 ± 0.9	−0.3 ± 0.3	−0.7 ± 0.8
			Water 317			
Asp145 (OD1)	95.6	2.8 ± 0.3	15.1 ± 10.6	−12.8 ± 4.5	1.7 ± 1.6	−11.1 ± 3.7
Asp145 (N-H)	31.6	3.4 ± 0.3	28.1 ± 10.2			
Lys33 (NZ-HZ)	88.0	3.3 ± 0.3	31.1 ± 15.6	−5.0 ± 5.9	0.3 ± 1.0	−4.7 ± 5.3
			Water 319			
Asp127 (OD2)	52.9	2.9 ± 0.4	19.4 ± 15.3	−6.8 ± 5.9	0.4 ± 1.3	−6.4 ± 4.9
Tyr15 (OH-HH)	51.8	2.9 ± 0.3	16.5 ± 13.9	−2.4 ± 3.7	0.4 ± 1.1	−2.0 ± 2.9
Asp145 (O)	43.1	2.7 ± 0.2	11.9 ± 7.6	−4.5 ± 4.1	0.5 ± 1.0	−4.0 ± 3.4
Ala149 (N-H)	27.6	3.4 ± 0.3	25.9 ± 15.4	−1.4 ± 1.4	−0.2 ± 0.4	−1.6 ± 2.2
			Water 320			
Asp127 (OD2)	97.1	2.7 ± 0.1	9.9 ± 6.2	−14.0 ± 4.0	1.5 ± 1.4	−12.5 ± 3.5
Asp127 (O)	33.5	3.6 ± 0.2	46.8 ± 8.8			
His125 (NE2)	94.6	2.9 ± 0.2	15.7 ± 8.8	−5.0 ± 2.0	0.7 ± 1.0	−4.3 ± 1.6
Asn132 (ND2-H)	37.6	3.4 ± 0.3	12.9 ± 7.4	−1.6 ± 0.6	−0.2 ± 0.1	−1.4 ± 0.6
Tyr15 (OH-HH)	25.5	3.1 ± 0.5	24.7 ± 15.9	−2.2 ± 2.6	0.1 ± 0.9	−2.1 ± 1.9
			Water 323			
Asp145 (N-H)	50.5	3.4 ± 0.4	37.0 ± 12.2	−3.9 ± 7.4	0.7 ± 1.5	−3.2 ± 6.2
Asp145 (OD1)	40.6	2.7 ± 0.1	13.9 ± 8.0			
			Water 348			
Leu55 (O)	99.9	2.8 ± 0.1	14.8 ± 8.9	−6.4 ± 1.6	0.8 ± 1.1	−5.6 ± 0.9
Asn59 (OD1)	67.4	2.8 ± 0.2	13.9 ± 8.5	−4.4 ± 3.3	0.5 ± 1.0	−3.9 ± 2.6
Leu66 (N-H)	51.8	3.7 ± 0.2	30.7 ± 9.6	−0.5 ± 0.2	−0.3 ± 0.1	−0.8 ± 0.3
			Water 363			
Ala149 (N-H)	64.4	3.4 ± 0.3	26.0 ± 13.5	−1.8 ± 1.5	−0.2 ± 0.1	−2.0 ± 1.5
Asp145 (O)	60.8	2.9 ± 0.4	19.2 ± 16.1	−5.1 ± 4.0	0.6 ± 1.1	−4.5 ± 3.2
Gly147 (N-H)	38.2	3.3 ± 0.3	42.0 ± 8.5	−3.1 ± 2.1	0.7 ± 0.9	−2.4 ± 1.5
His125 (O)	38.3	2.9 ± 0.2	17.3 ± 10.4	−3.7 ± 2.1	−0.4 ± 0.2	−3.3 ± 1.6
His125 (NE2)	26.3	3.4 ± 0.3	39.3 ± 15.4			
			Water 377			
Asp145 (OD2)	76.3	2.7 ± 0.2	12.1 ± 10.6	−12.6 ± 4.9	1.5 ± 1.5	−11.1 ± 4.1
Thr14 (OG1-HG1)	57.1	2.9 ± 0.3	15.8 ± 13.8	−4.3 ± 3.7	0.7 ± 1.3	−3.6 ± 2.9
			Water 394			
Gly147 (O)	81.9	2.8 ± 0.2	13.0 ± 8.6	−4.1 ± 2.6	0.8 ± 1.0	−3.3 ± 2.1
Ala151 (N-H)	65.1	3.3 ± 0.3	34.2 ± 12.4	−1.4 ± 1.4	−0.1 ± 0.1	−1.3 ± 1.2
			Water 403			
Val64 (O)	100.0	2.7 ± 0.1	11.3 ± 6.2	−7.6 ± 1.2	1.3 ± 1.1	−6.3 ± 0.7
			Water 412			
Phe146 (N-H)	98.2	3.3 ± 0.3	42.0 ± 8.5	−1.6 ± 0.2	−0.3 ± 0.1	−1.9 ± 0.2
Leu143 (O)	98.0	2.9 ± 0.2	16.4 ± 9.7	−5.9 ± 1.6	0.5 ± 0.9	−5.4 ± 1.0
			Water 418			
Lys34 (O)	100.0	2.8 ± 0.2	14.7 ± 9.1	−4.3 ± 1.5	0.8 ± 1.0	−3.5 ± 1.7
Val17 (O)	71.8	3.1 ± 0.3	22.8 ± 14.4	−3.2 ± 2.3	−1.1 ± 0.1	−4.3 ± 2.0
Gly16 (O)	20.1	3.1 ± 0.3	28.9 ± 15.3	−0.6 ± 0.3	−0.4 ± 0.2	−1.0 ± 0.4
			Water 427			
Rhe146 (O)	100.0	2.8 ± 0.2	16.9 ± 10.1	−4.9 ± 1.4	0.7 ± 1.1	−4.2 ± 1.2
			Water 6690			
Leu83 (O)	79.4	2.8 ± 0.2	16.8 ± 10.6	−5.4 ± 3.3	0.6 ± 1.1	−4.8 ± 2.8
Leu83 (N-H)	57.1	3.3 ± 0.3	24.9 ± 12.5			
			Water 6795			
Asn132 (ND2-HD22)	62.0	3.0 ± 0.2	14.3 ± 9.4	−3.1 ± 3.3	−0.3 ± 0.2	−3.4 ± 2.8
Thr14 (OG1-HG1)	33.1	3.5 ± 0.4	42.0 ± 13.8	−0.6 ± 0.8	−0.1 ± 0.6	−0.7 ± 0.7
Thr14 (OG1)	21.4	3.6 ± 0.3	41.3 ± 12.7			

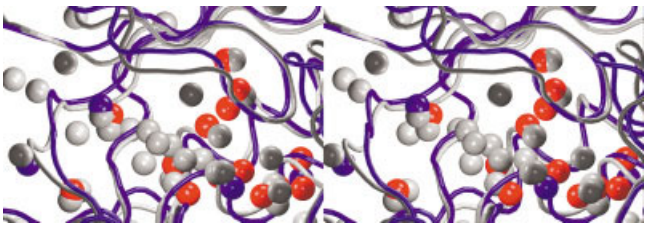


Fig. 5. Comparison of the locations of crystallographic and MD structural water molecules for free CDK2 (X-ray, gray; X-ray structural water molecules in MD simulation, red; structural water molecules from bulk, blue).

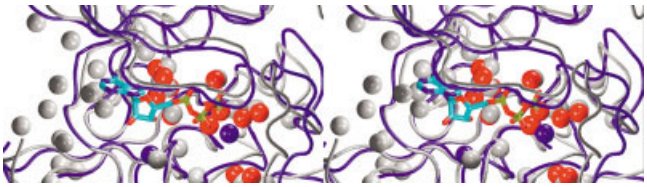


Fig. 8. Comparison of the locations of crystallographic and MD structural water molecules for CDK2/ATP complex (X-ray, gray; X-ray structural water molecules in MD simulation, red; structural water molecules from bulk, blue).

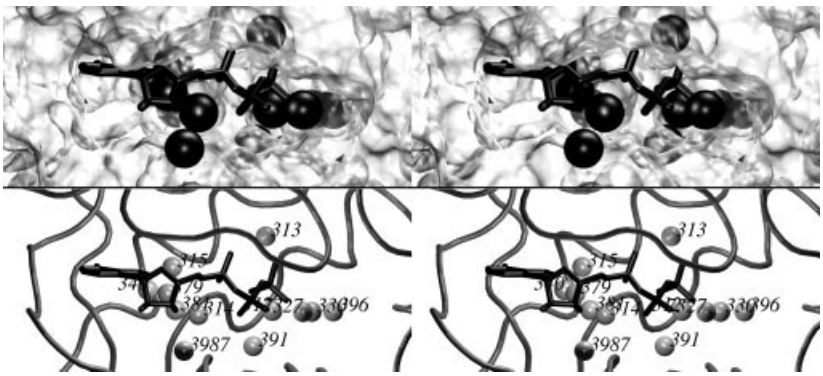


Fig. 7. The location of structural water molecules in the CDK2 active site found for CDK2/ATP complex. ATP is represented by the tube model, the active site of the enzyme by the solvent accessible surface, and water molecules by the CPK model (**top**). For better clarity, the same situation is pictured without the surface, and the positions of water molecules are represented by oxygen atoms (balls; **bottom**). Water molecules numbered as 400 or lower are those detected by X-ray crystallography.

with the ASP 145 side-chain. Moreover, these two water molecules weakly interact with the nitrogen N7 of the purine ring, and water 314 also has short contact with O3* oxygen of the ATP ribose. The ATP ribose is also weakly solvated by water molecule 3987, which creates a hydrogen bond with the Asp86 side-chain. Water 313 creates a bridge between the Tyr15 and Gly16 backbones, and the phosphate group of the ATP. Similar behavior is exhibited by water 312, which interacts with the Asp127 and Asn132 side-chains and the gamma phosphate group of the ATP. Water 330 solvates the Asp127 side-chain and for a significant part of the trajectory creates bridges between the Thr14 and Lys129 side-chains and the gamma phosphate group of the ATP. The calculated strong repulsion of water 330 with the Lys129 side-chain is balanced by attractive interactions with other residues. Water molecule 391 contributes to the tertiary structure stabilization of the CDK2 due to the bridge between the Ala144 backbone and the His125 side-chain. Similar behavior is seen for water 396, which creates a bridge between the Tyr15 backbone and the His125 side-chain. High interaction energies are caused by interactions of charged phosphate group with water molecules. Water 3987 exhibits strong interaction with the Asp86 side-chain. The H-bond between the water and Asp86 lasts only for 17% of the trajectory, but the interaction energy is large, with large deviations due to interaction with charged residue. This water molecule also exhibits repulsive interaction with the ATP and a short interaction with the Lys89 side-chain. The large interaction energy deviations indicate the high mobility of this water. Structural water 7327 contributes to the stability of the ATP phosphate side-chain and Thr14 side-chain. High interaction energies are caused by strong interactions between charged atoms of the ATP and the oxygens of the water molecules.

CDK2/Roscovitine Complex

Water molecules location. In this case, the water molecules create a large cluster stabilized by a network of H-bonds, which is remarkably more stable than the clusters in the other simulations. The cluster of 8 water molecules is positioned between the LYS 33 side-chain and polar phosphate pocket of the CDK2 active site (residues Thr14, Tyr15, Lys129, Asn132, and Asp145). The situation in the active site is shown in Figure 9. One water molecule (881) has a strong contact with the phenylene ring of the roscovitine. The calculated interaction energy of the water molecule is -4.9 ± 1.4 kcal/mol. The standard AMBER H-bond analysis did not indicate any strong hydrogen bond between this water molecule and roscovitine atoms. This is because the interaction with the aromatic ring can be considered as a X—H…π-hydrogen bond.[40] Water molecule 881 has a strong hydrogen bond with the Asp86 residue. This water molecule creates a bridge between the roscovitine phenylene ring and Asp86. The isopropyl substituent on the N9 roscovitine atom induces a conformational change of the Lys33 side-chain. The change opens a deep cavity from the purine pocket through the whole

TABLE III. Interaction Energies and H-Bond Analyses Calculated for MD Structural Water Molecules in the CDK2/ATP Trajectory

Residue	Occupancy [%]	Distance [Å]	H-bond angle [degrees]	Interaction energy [kcal/mol]		
				Electrostatic	VdW	Total
			Water 312			
Asp127 (OD)	38.7	2.8 ± 0.3	14.1 ± 11.6	−2.1 ± 1.7	−0.7 ± 0.1	−2.3 ± 1.7
Asn132 (ND-HD)	91.5	3.0 ± 0.3	23.8 ± 10.9	−1.8 ± 1.4	−1.8 ± 0.8	−2.3 ± 1.3
ATP (OG)	100.0	2.9 ± 0.3	24.4 ± 14.8	−24.7 ± 5.3	1.8 ± 2.0	−23.6 ± 5.0
			Water 313			
Tyr15 (N-H)	98.0	3.1 ± 0.2	22.8 ± 10.6	−2.7 ± 1.5	−0.1 ± 0.5	−2.8 ± 1.3
Gly6 (N-H)	99.3	3.0 ± 0.1	20.5 ± 9.4	−3.1 ± 1.1	0.4 ± 0.8	−2.7 ± 0.9
ATP (OA)	99.8	2.7 ± 0.1	13.6 ± 6.3	−31.9 ± 3.1	3.8 ± 1.9	−28.5 ± 2.1
ATP (OG)	100.0	2.6 ± 0.1	8.9 ± 5.1			
			Water 314			
Gln131 (O)	99.0	2.7 ± 0.2	14.4 ± 8.1	−6.8 ± 1.6	1.4 ± 1.4	−5.4 ± 1.4
			Water 315			
Asp145 (OD1)	100.0	2.8 ± 0.2	13.3 ± 8.3	−2.1 ± 0.4	−0.1 ± 0.1	−2.1 ± 0.3
			Water 330			
Thr14 (OG1)	74.0	3.2 ± 0.3	31.2 ± 12.7	−0.9 ± 1.1	−0.1 ± 0.3	−1.1 ± 1.1
Asp127 (OD)	100.0	2.7 ± 0.2	13.4 ± 9.0	−10.4 ± 9.2	1.5 ± 1.6	−8.8 ± 8.1
Lys129 (NZ-HZ)	58.6	3.3 ± 0.4	29.8 ± 14.2	3.9 ± 8.0	0.1 ± 0.1	3.9 ± 7.5
ATP (OG)	60.5	2.7 ± 0.2	16.6 ± 9.3	−15.4 ± 10.3	0.7 ± 1.3	−14.6 ± 9.5
			Water 340			
Leu55 (O)	69.2	3.1 ± 0.3	18.3 ± 9.9	−5.0 ± 2.2	0.2 ± 1.1	−4.8 ± 1.5
Asn59 (OD1)	48.5	3.0 ± 0.3	27.3 ± 14.7	−4.9 ± 2.1	0.2 ± 0.9	−4.7 ± 1.5
Val64 (O)	57.3	2.9 ± 0.2	17.1 ± 10.8	−2.8 ± 3.7	0.1 ± 0.2	−2.6 ± 3.2
Leu66 (N-H)	58.5	3.4 ± 0.4	27.2 ± 12.8	−0.6 ± 1.6	−0.3 ± 0.5	−0.9 ± 1.3
			Water 379			
Asp145 (OD)	48.8	3.2 ± 0.6	21.9 ± 15.6	−1.4 ± 5.4	−0.1 ± 0.1	−1.4 ± 4.9
ATP (N7)	69.5	2.8 ± 0.1	14.3 ± 7.5	−5.3 ± 3.8	1.7 ± 1.7	−3.6 ± 2.4
			Water 381			
Val64 (O)	20.2	3.0 ± 0.3	22.7 ± 12.0	−3.1 ± 2.3	−0.1 ± 0.1	−3.1 ± 1.9
Phe146 (N-H)	64.7	3.3 ± 0.3	25.6 ± 11.6	−1.9 ± 1.3	−0.3 ± 0.2	−2.2 ± 1.3
			Water 391			
Ala144 (O)	100.0	2.8 ± 0.2	13.7 ± 8.1	−5.7 ± 1.5	1.0 ± 1.0	−4.7 ± 0.9
His125 (ND-HD)	100.0	2.9 ± 0.1	19.2 ± 7.6	−6.4 ± 1.5	1.1 ± 1.1	−5.3 ± 1.1
			Water 396			
Tyr15 (N-H)	97.7	2.9 ± 0.2	10.0 ± 7.2	−5.8 ± 2.4	1.3 ± 1.4	−4.5 ± 1.6
His125 (NE)	79.7	3.1 ± 0.3	20.5 ± 12.3	−4.7 ± 3.8	−0.2 ± 0.7	−4.9 ± 3.4
			Water 3987			
Lys89 (NZ-H)	20.0	3.0 ± 0.2	21.9 ± 10.5	−3.7 ± 0.5	0.0 ± 0.1	−3.7 ± 0.5
ATP (O2*-H)	18.7	3.7 ± 0.2	42.8 ± 5.2	1.4 ± 3.6	−0.4 ± 0.2	1.0 ± 3.6
ATP (O2*)	11.7	3.7 ± 0.2	34.1 ± 11.5			
Asp86 (OD)	16.7	2.9 ± 0.3	19.9 ± 12.5	−11.8 ± 7.9	1.1 ± 1.5	−10.7 ± 7.2
			Water 7327			
Thr14 (OG)	82.0	2.9 ± 0.2	15.7 ± 8.7	−1.2 ± 1.8	0.4 ± 0.9	−0.8 ± 1.5
ATP (OB)	100.0	2.6 ± 0.1	8.8 ± 5.1	−21.7 ± 5.4	2.4 ± 1.8	−19.4 ± 4.8
ATP (OG)	29.7	3.7 ± 0.2	42.5 ± 6.1			

enzyme to the bulk. The cavity entrance is created by hydrophobic amino acids (Val64, Phe80, Ala144, and Leu148). However, the interior of the cavity is hydrophilic and it is filled by 5 water molecules (332, 336, 341, 344, and 357). These water molecules interact with polar groups of the Leu55, Asp145, and Phe146 backbone. Although water molecules 332, 336, and 341 are positioned far from the roscovitine molecule, we have considered them in our analyses, because these water molecules show large fluctuations during the simulation involving the roscovitine molecule. In the cases of the free CDK2 and CDK2/IPA complex MD simulations, the cavity is smaller. An exchange of the positions of water molecules in the cavity and the penetration of the bulk water molecules behind the roscovitine into the cavity was observed. The exchange time of the positions of water molecules in the cavity is from 200 to 300 ps. The comparison of the positions of structural water molecules detected by MD simulations with crystallographic water molecules is shown in Figure 10. Major changes in water molecule location in comparison with other trajectories are visible in the purine base pocket. The cluster in the phosphate pocket of the CDK2 active site is reorganized to create a configuration with a larger number of hydrogen bonds between water mol-

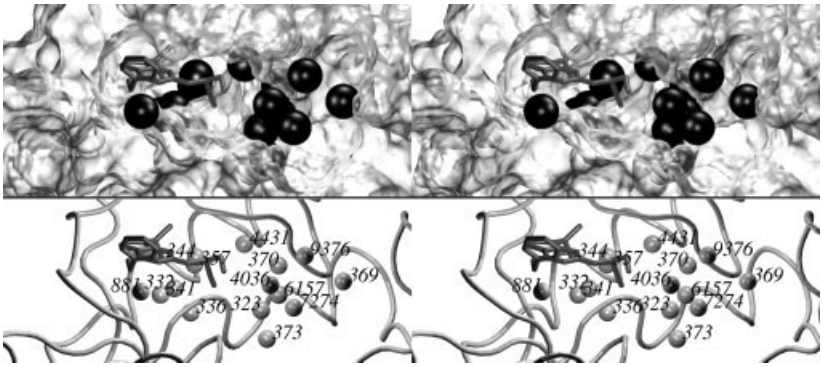


Fig. 9. The location of structural water molecules in the enzyme active site found for CDK2/roscovitine complex. Roscovitine is represented by the tube model, active site of the CDK2 is modeled by the solvent-accessible surface, the water molecules by the CPK model (**top**). For better clarity, the same situation is pictured without the surface, and the positions of water molecules are represented by oxygen atoms (balls; **bottom**). Water molecules numbered as 394 or lower are those detected also by X-ray crystallography.

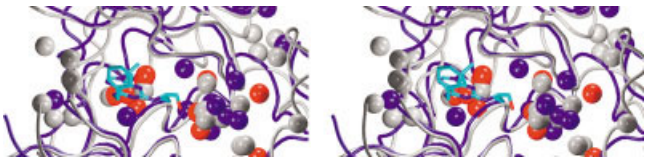


Fig. 10. Comparison of the locations of crystallographic and MD structural water molecules for the CDK2/roscovitine complex (X-ray structure, gray; X-ray structural water molecules in MD simulation, red; structural water from bulk, blue).

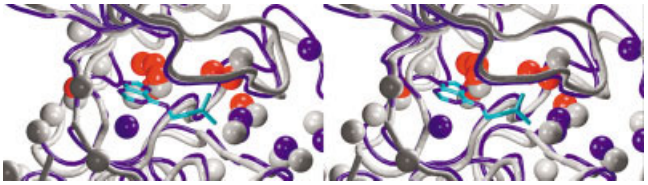


Fig. 12. Comparison of the locations of crystallographic and MD structural water molecules for the CDK2/isopentenyladenine complex (X-ray structure, gray; X-ray structural water molecules in MD simulation, red; structural water from bulk, blue).

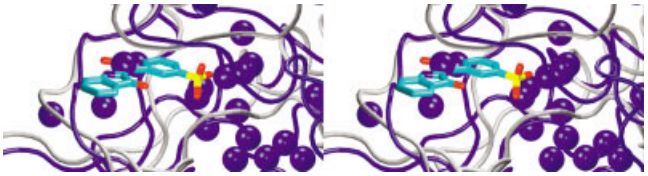


Fig. 13. Superimposition of structural water molecules found in free CDK2 MD simulations (blue) with the polar sulphonate group of CDK2/indirubin-5-sulphonate complex (color). [X-ray obtained from the Brookhaven protein crystallographic database (PDB code: 1e9h).]

COLOR

Fig. 14. Superimposition of structural water molecules found in free CDK2 MD simulations (blue) with the polar group of CDK2/hymenialdisine complex (color). [X-ray obtained from the Brookhaven protein crystallographic database (PDB code: 1dm2).]

COLOR

ecules and the enzyme. Partial exchange of X-ray water molecules by bulk MD waters in the phosphate part of the CDK2 active site is observed.

Interaction energies. The interaction energies calculated from MD simulations between water molecules and amino acids in the CDK2 active site are summarized in Table IV.

As in the previous case, the highest interaction energies generate interactions between charged side-chains of polar amino acids and stable water molecules. The residues in the active site interacting with only one structural water molecule are Asp145, Phe146, and Thr14.

Water molecule 323 creates a bridge between the Asp145 side-chain and the Lys129 or Asp127 side-chains. Water molecule 341 stabilizes interaction between the Leu55 and Asn59 side-chains. The large number of interactions found for water molecules 370, 373, and 4036 indicates the positional fluctuations of these water molecules.

During the relaxation phase of the MD simulation, some bulk water molecules entered the active site and remained stable there for the rest of the trajectory. This is the case with water molecules 4431 and 7274. Water 4431 stabilizes the position of the Lys33 side-chain with respect to the Glu12 and Gly16 backbone. Water 7274 hydrates the Thr14 backbone and exhibits also a short interaction with the Lys129 side-chain. The cluster of water molecules in the phosphate part of the active site is enlarged by 3 other water molecules (4036, 6157, and 9376). These water molecules joined the cluster during the production phase of the simulation. This is reflected by high deviations in interaction energies and in the number of peptide residues that are in contact with these water molecules, and also by low H-bond occupancy of observed hydrogen bonds.

CDK2/Isopentenyladenine Complex

Water molecules location. The configuration of 11 structural water molecules in the active site found in the CDK2/isopentenyladenine MD simulation is depicted in the Figure 11.

One water molecule interacting with the purine ring (water 7067) was found at a similar position as water 881 in the case of CDK2/roscovitine complex. The orientation

TABLE IV. Interaction Energies and H-Bond Analyses Calculated for MD Structural Water Molecules in the CDK2/Roscovitine Trajectory

Residue	Occupancy [%]	Distance [Å]	H-bond angle [degrees]	Interaction energy [kcal/mol]		
				Electrostatic	VdW	Total
			Water 323			
Asp145 (OD2)	82.9	2.7 ± 0.1	10.7 ± 6.3	−10.3 ± 7.0	1.3 ± 1.6	−9.0 ± 6.0
Lys129 (NZ-HZ)	49.3	3.3 ± 0.3	40.2 ± 10.8	−4.9 ± 4.5	−0.2 ± 0.5	−5.1 ± 4.2
Asp127 (OD2)	33.3	2.9 ± 0.4	13.4 ± 13.8	−4.5 ± 7.2	0.3 ± 1.4	−4.2 ± 6.1
			Water 332			
Asp145 (N-H)	59.9	3.1 ± 0.2	23.3 ± 13.8	−1.7 ± 1.0	−0.3 ± 0.3	−2.0 ± 1.0
			Water 336			
Phe146 (O)	99.9	2.8 ± 0.2	18.0 ± 11.0	−5.9 ± 1.5	0.7 ± 1.1	−5.2 ± 1.0
			Water 341			
Asn59 (ND2-HD2)	75.8	3.0 ± 0.2	15.0 ± 9.1	−6.3 ± 2.3	0.6 ± 1.0	−5.7 ± 1.8
Asn59 (OD1)	18.7	2.8 ± 0.2	18.5 ± 11.6			
Leu55 (O)	73.5	2.9 ± 0.2	13.8 ± 9.0	−3.9 ± 2.5	0.4 ± 0.9	−3.5 ± 2.1
			Water 344			
Asp145 (OD1)	74.1	2.6 ± 0.1	10.3 ± 6.3	−9.5 ± 6.9	1.9 ± 1.9	−7.6 ± 5.6
Asp145 (N-H)	45.8	3.5 ± 0.4	32.1 ± 11.4			
Lys33 (NZ-HZ)	54.2	3.2 ± 0.4	28.5 ± 12.6	−3.0 ± 4.4	0.1 ± 0.8	−2.9 ± 4.0
			Water 357			
Asp145 (N-H)	37.1	3.2 ± 0.4	28.7 ± 15.7	−1.5 ± 1.3	−0.2 ± 0.2	−1.7 ± 1.3
Val64 (O)	32.6	3.0 ± 0.3	15.9 ± 10.8	−2.0 ± 3.0	−0.2 ± 0.2	−2.2 ± 2.6
			Water 369			
Thr14 (N-H)	57.8	3.1 ± 0.2	15.5 ± 7.5	−1.9 ± 2.1	0.1 ± 0.1	−1.8 ± 1.5
			Water 370			
Tyr15 (N-H)	100.0	3.0 ± 0.2	12.8 ± 7.2	−3.9 ± 0.9	0.3 ± 0.9	−3.6 ± 0.8
Asp145 (OD2)	99.5	2.7 ± 0.1	13.3 ± 7.7	−14.2 ± 3.1	2.2 ± 1.5	−12.0 ± 2.3
Lys33 (NZ-HZ)	91.8	3.0 ± 0.2	24.3 ± 13.2	−4.3 ± 5.2	0.6 ± 0.9	−3.7 ± 4.8
Thr14 (OG1)	44.4	3.1 ± 0.3	13.2 ± 8.9	−1.7 ± 1.8	−0.1 ± 0.7	−1.8 ± 1.4
Thr14 (N-H)	21.1	3.1 ± 0.2	50.6 ± 6.0			
Gly16 (N-H)	40.6	3.7 ± 0.2	34.4 ± 10.7	0.3 ± 0.8	−0.2 ± 0.1	0.1 ± 0.8
			Water 373			
Asp127 (OD2)	100.0	2.7 ± 0.1	11.6 ± 7.1	−16.6 ± 6.7	2.3 ± 2.0	−14.3 ± 5.4
Asp127 (O)	85.5	2.9 ± 0.2	26.9 ± 11.3			
His125 (ND1-HD1)	53.9	3.4 ± 0.4	40.8 ± 14.6	−0.2 ± 0.2	−0.1 ± 0.1	−0.3 ± 0.2
Lys129 (NZ-HZ)	26.2	3.3 ± 0.3	32.6 ± 17.8	−0.3 ± 0.4	−0.2 ± 0.5	−0.5 ± 0.5
			Water 881			
Asp86 (OD2)	87.6	2.7 ± 0.1	10.6 ± 6.3	−13.5 ± 4.8	1.7 ± 1.6	−11.9 ± 4.0
Asp86 (N-H)	56.4	3.4 ± 0.3	25.9 ± 11.2			
Gln85 (NE2-HE22)	21.9	3.4 ± 0.4	34.3 ± 18.3	−2.0 ± 2.9	−0.2 ± 0.5	−2.2 ± 2.7
			Water 4036			
Gly13 (O)	41.6	2.9 ± 0.3	19.2 ± 11.5	−2.0 ± 2.5	0.2 ± 0.7	−1.8 ± 2.1
Gly13 (N-H)	20.0	3.6 ± 0.2	36.9 ± 12.1			
Thr158 (OG1-HG1)	27.4	3.3 ± 0.5	30.6 ± 17.0	−1.5 ± 2.7	0.1 ± 0.8	−1.4 ± 2.1
Tyr159 (OH)	27.3	3.5 ± 0.3	42.7 ± 11.4	−0.5 ± 0.7	−0.3 ± 0.4	−0.8 ± 0.8
			Water 4431			
Lys33 (NZ-HZ)	91.3	2.9 ± 0.2	23.5 ± 12.3	−12.3 ± 5.7	1.1 ± 1.3	−11.2 ± 5.0
Glu12 (O)	74.8	3.0 ± 0.3	38.1 ± 15.9	−4.4 ± 3.2	−0.1 ± 0.4	−4.5 ± 2.9
Gly16 (O)	55.5	2.8 ± 0.2	13.5 ± 8.9	−3.8 ± 3.2	0.8 ± 1.2	−3.0 ± 2.5
			Water 6157			
Lys129 (NZ-HZ)	37.8	3.0 ± 0.3	28.5 ± 12.5	−3.8 ± 5.6	0.2 ± 0.8	−3.6 ± 5.1
Thr14 (N-H)	26.0	3.2 ± 0.2	17.8 ± 6.9	−1.0 ± 1.2	−0.1 ± 0.3	−1.1 ± 1.2
Lys33 (NZ-HZ)	15.2	2.8 ± 0.1	33.3 ± 12.4	−2.1 ± 5.5	0.2 ± 0.7	−1.9 ± 5.0
Gly16 (O)	10.4	2.7 ± 0.1	12.3 ± 7.2	−0.9 ± 2.3	0.2 ± 0.7	−0.7 ± 1.7
Glu12 (O)	10.0	3.1 ± 0.4	27.7 ± 12.0	−1.1 ± 2.0	0.0 ± 0.3	−1.1 ± 1.9
			Water 7274			
Thr14 (O)	48.9	2.8 ± 0.2	12.8 ± 8.4	−3.3 ± 2.9	0.4 ± 1.0	−2.9 ± 2.3
Thr14 (N-H)	16.7	3.1 ± 0.2	16.9 ± 9.6			
Lys129 (NZ-HZ)	17.0	3.1 ± 0.4	33.1 ± 13.4	−2.0 ± 3.9	0.0 ± 0.4	−2.0 ± 3.7
			Water 9376			
Asp86 (OD1)	29.6	2.8 ± 0.3	13.7 ± 9.9	−2.6 ± 5.8	0.2 ± 0.8	−2.4 ± 5.2
Asp145 (OD2)	15.2	2.8 ± 0.3	15.7 ± 12.4	−2.1 ± 5.0	0.3 ± 0.8	−1.8 ± 4.4

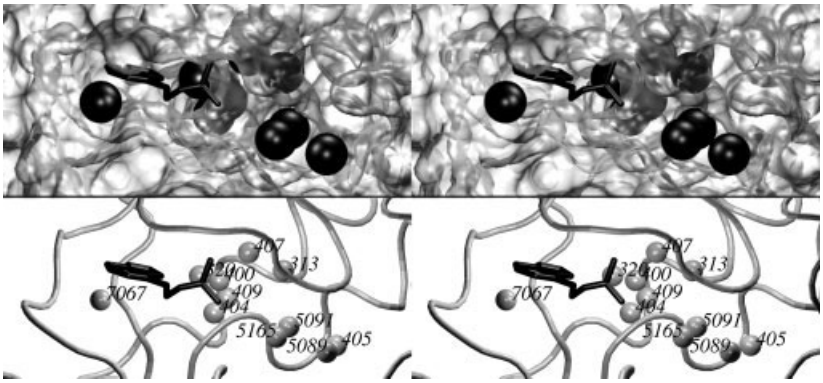


Fig. 11. The location of structural water molecules in the enzyme active site found for CDK2/isopentenyladenine complex. Isopentenyladenine is represented by the tube model; active site of the CDK2 is modeled by the solvent accessible surface, water molecules by the CPK model (**top**). For better clarity, the same situation is pictured without the surface, and the positions of water molecules are represented by oxygen atoms (balls; **bottom**). Water molecules numbered as 411 or lower are those detected also by X-ray crystallography.

of the isopentenyladenine purine ring in the CDK2 active site differs from the orientation of the roscovitine purine ring and is more favorable for interactions with water molecules. The hydrophobic isopentenyladenine side-chain interacts with only 1 water molecule (320). The cavity formed by Val64, Phe80, Ala144, and Leu148 behind the purine pocket is smaller than in the case of CDK2/roscovitine, and it is filled by 4 water molecules (320, 400, 404, and 409). Water molecules, as in the previous cases, also hydrate the Asp145 side-chain.

The phosphate pocket of the active site is filled by just 4 water molecules. The smaller number of water molecules could be explained by the influence of the hydrophobic side-chain of isopentenyladenine. The comparison of the positions of the structural water molecules found by MD simulation with crystallographic water molecules is shown in Figure 12.

A small reorientation of water molecules is seen in the phosphate pocket. In this case, bulk water molecules replaced three X-ray structural water molecules.

Interaction energies. Calculated interaction energies are summarized in Table V.

As in the previous cases, 1 water molecule (313) stabilizes the position of the Asp145 side-chain relative to the backbone of the Tyr15 or Gly16. Water 320 arranges the interaction between the Asp145, Lys side-chains and the C6 isopentenyladenine side-chain. The interaction analysis of water molecules 400, 404 and 405 indicates the positional exchange of these waters in the active site. Water 405 stabilizes the orientation of the Asp145 side-chain to the Gly16 backbone. Similar stabilization is seen in the case of water molecule 409 (Asn59 side-chain to the Leu55 backbone). Bulk water molecule 5089 stabilizes the position of the Thr14, Lys129, and Asn132 side-chains. Water molecule 7067 creates a bridge between the Leu83 backbone and the isopentenyladenine molecule. There are also some water molecules (5089, 5091, 5165) found that create a cluster with some of the above-mentioned water molecules and stabilize the entire system.

The statistical analysis of binding energy data introduced in the above tables shows that the correlation coefficients between the occupancies of H-bond and interaction energies as calculated by the force field are between 0.4 and 0.7, which indicates correlation on 95% probability level. We have also calculated the correlation between energies and H-bond distances or H-bond angles. The calculated correlation coefficients between the energies and H-bond distances are in range of 0.6 to 0.7. It also indicates correlation on 95% probability level. The correlation between the energies and H-bond angles is not so significant (calculated between 0.3 and 0.6). It is caused by predomination of electrostatic term in interaction energy calculation.

Comparison With Other X-Ray Structures From the PDB Database

The positions of MD structural water molecules in the free CDK2 active site were compared with the positions of polar groups of known nonpeptidic inhibitors. For this, we used X-ray data of 17 complexes of CDK2 with nonpeptidic inhibitors stored in the Brookhaven PDB. As expected, water molecules interacting with the Glu81 and Leu83 backbone found in MD were replaced by a polar group of inhibitors. These H-bond interactions are important for substrate binding to the active site.[41] Inhibitors that have a polar group oriented to the ribose pocket have this group oriented in the same direction as structural water molecules that were found. This is illustrated in Figures 13 and 14 for the inhibitors hymenialdisine and indirubin-5-sulphonate, respectively.

CONCLUSIONS

We have presented a very detailed analysis of solvent behavior over 4 MD trajectories on CDK2 and its substrate/inhibitor complexes. The data obtained were compared with X-ray information. The presented method is complementary to that of X-ray crystallography. The static picture of structure from X-ray data is completed by information about the dynamics of molecules. The MD method also provides information about interaction energies between

TABLE V. Interaction Energies and H-Bond Analyses Calculated for MD Structural Water Molecules in the CDK2/Isopentenyladenine Trajectory

Residue	Occupancy [%]	Distance [Å]	H-bond angle [degrees]	Interaction energy [kcal/mol]		
				Electrostatic	VdW	Total
			Water 313			
Asp145 (OD1)	97.8	2.8 ± 0.2	12.4 ± 9.2	−13.4 ± 3.7	1.2 ± 1.3	−12.2 ± 3.0
Tyr15 (N-H)	77.4	3.1 ± 0.2	16.0 ± 7.4	−2.5 ± 1.2	−0.1 ± 0.7	−2.6 ± 1.5
Gly16 (N-H)	49.9	3.5 ± 0.3	39.5 ± 10.8	0.1 ± 0.9	−0.3 ± 0.2	−0.2 ± 0.8
			Water 320			
Asp145 (OD1)	100.0	2.8 ± 0.4	17.7 ± 13.2	−5.8 ± 7.5	0.9 ± 1.6	−4.9 ± 6.3
Lys33 (NZ-H)	93.7	3.0 ± 0.2	28.3 ± 15.3	−8.3 ± 5.6	0.8 ± 1.1	−7.5 ± 5.0
Ipa299 (N6-H)	35.8	3.3 ± 0.3	16.7 ± 7.9	−0.7 ± 2.0	−0.2 ± 0.6	−0.9 ± 1.7
			Water 400			
Asp145 (OD2)	54.1	2.7 ± 0.1	13.0 ± 7.4	−6.3 ± 6.3	1.0 ± 1.5	−5.3 ± 5.3
Lys33 (NZ-H)	25.2	3.1 ± 0.3	38.1 ± 13.5	−3.5 ± 4.4	0.2 ± 0.7	−3.3 ± 4.0
Asn59 (ND-H)	22.8	3.2 ± 0.3	20.5 ± 13.5	−1.6 ± 2.9	0.2 ± 0.7	−1.4 ± 2.6
			Water 404			
Phe146 (O)	60.1	2.9 ± 0.3	12.7 ± 7.6	−3.9 ± 2.9	0.1 ± 0.1	−3.8 ± 2.6
Asp145 (O)	41.7	3.2 ± 0.3	28.0 ± 16.0	−2.6 ± 0.9	0.1 ± 0.1	−2.5 ± 0.9
			Water 405			
Gly147 (N-H)	73.4	3.2 ± 0.3	34.8 ± 12.2	−6.2 ± 4.2	−0.4 ± 0.1	−6.6 ± 3.5
Leu148 (N-H)	31.1	3.3 ± 0.3	50.1 ± 6.1	−2.2 ± 1.4	−0.4 ± 0.2	−2.6 ± 1.2
Tyr15 (OH)	31.0	3.1 ± 0.3	24.1 ± 14.2	−2.5 ± 3.1	0.4 ± 1.2	−2.1 ± 2.2
His125 (NE2)	26.0	3.3 ± 0.3	32.7 ± 16.5	−2.7 ± 2.1	−0.2 ± 0.8	−2.9 ± 1.6
			Water 407			
Gly16 (O)	78.1	2.9 ± 0.2	16.7 ± 11.1	−4.0 ± 2.7	0.6 ± 1.1	−3.5 ± 2.1
Asp145 (OD2)	51.0	3.2 ± 0.5	32.5 ± 18.5	−1.5 ± 6.9	0.1 ± 0.8	−1.4 ± 6.3
			Water 409			
Asn59 (OD1)	100.0	2.8 ± 0.2	17.2 ± 10.5	−5.8 ± 2.6	0.6 ± 1.1	−5.2 ± 2.0
Leu55 (O)	75.9	3.0 ± 0.3	18.2 ± 11.3	−3.7 ± 2.5	0.1 ± 0.8	−3.6 ± 2.1
			Water 5089			
Lys129 (NZ-H)	78.8	3.0 ± 0.2	28.8 ± 13.9	−11.3 ± 5.1	0.6 ± 1.2	−10.7 ± 4.3
Thr14 (OG1)	58.8	2.9 ± 0.3	20.9 ± 14.2	−2.1 ± 2.7	0.2 ± 0.8	−1.9 ± 2.1
Asn132 (ND2-H)	39.3	3.1 ± 0.2	20.5 ± 11.3	−2.1 ± 2.6	0.1 ± 0.6	−2.0 ± 2.3
			Water 5091			
Thr14 (N-H)	28.8	3.2 ± 0.3	24.7 ± 10.8	−0.6 ± 1.2	−0.1 ± 0.4	−0.7 ± 1.0
Asn132 (ND2-H)	27.9	3.1 ± 0.4	19.7 ± 11.6	−1.4 ± 2.7	−0.1 ± 0.7	−1.5 ± 2.3
Asp145 (OD2)	26.5	2.8 ± 0.2	12.3 ± 8.5	−5.2 ± 7.8	0.6 ± 1.2	−4.6 ± 7.0
			Water 5165			
Thr14 (OG1-H)	46.3	3.3 ± 0.3	34.4 ± 12.3	−0.1 ± 2.0	−0.2 ± 0.6	−0.3 ± 1.6
Asn132 (ND2-H)	56.3	3.2 ± 0.3	33.4 ± 15.2	−1.8 ± 2.6	−0.1 ± 0.7	−1.9 ± 2.3
Asp127 (OD2)	35.9	2.8 ± 0.2	11.9 ± 9.2	−5.8 ± 7.5	0.6 ± 1.3	−5.2 ± 6.6
His125 (NE2)	32.9	3.0 ± 0.2	14.4 ± 8.5	−2.6 ± 3.0	0.2 ± 0.6	−2.4 ± 2.6
			Water 7067			
Ipa299 (N1)	92.9	3.1 ± 0.2	22.8 ± 12.2	−2.7 ± 1.6	−0.1 ± 0.7	−2.8 ± 1.3
Leu83 (O)	88.2	2.9 ± 0.3	16.5 ± 10.3	−4.5 ± 2.3	0.3 ± 0.8	−4.2 ± 1.9

molecules and is also helpful in the case of uncompleted X-ray data. In our case, the MD simulation detected a larger number of stable water molecules in the active site than X-ray crystallography. Moreover, it provided us with information about changes in the configuration of water molecules in the CDK2 active site.

The behavior of water molecules interacting with amino acids in the enzyme active site, and their interaction energies, provides information that is useful for the rational drug design of new potent and selective inhibitors. The structural water molecules found for free CDK2 were in the same positions as the nitrogens of the purine ring of the ATP and inhibitors. Appropriate replacement of the aforementioned water molecule in the active site leads to more selective inhibitors as demonstrated by the replacement of stable water molecules by polar groups of inhibitor.

The interaction of inhibitor with Asp86 was found for many compounds.[41] Therefore, one can suppose the replacing of the π-interacting water molecule by a hydrophilic group results in an increase in inhibitor activity. This assumption is demonstrated by higher activities of -OH substituted purine-like inhibitors.[42] The inclusion of the substrate into the active site evokes the conformational changes in this part, which allows the opening of the cavity through the enzyme and movement of the water molecules to the active site. The detailed analysis of solvent behavior and interaction energy information may be used as a tool

to design new inhibitors, because it indicates where the polar groups should be oriented.

ACKNOWLEDGMENTS

Our thanks to the Supercomputer Center Brno for providing us with computer time, and to Roger Turland for language corrections.

REFERENCES

1. Brooks CL III, Karplus M. Solvent effects on protein motion and protein effects on solvent motion. J Mol Biol 1989;208:159–181.
2. Thanki N, Thornton JM, Goodfellow JM. Influence of secondary structure on the hydration of serine, threonine and tyrosine residues in proteins. Protein Eng 1990;3:495–508.
3. Ni H, Sotriffer CA, McCammon JA. Ordered water and ligand mobility in the HIV-1 integrase-5CITEP complex: a molecular dynamics study. J Med Chem 2001;44:3043–3047.
4. García AE, Hummer G. Water penetration and escape in proteins. Proteins 2000;38:261–272.
5. Ladbury JE. Just add water!: the effect of water on the specificity of protein–ligand binding sites and its potential application to drug design. Chem Biol 1996;3:973–980.
6. Koellner G, Kryger G, Millard CB, Silman I, Sussman JL, Steiner T. Active site gorge and buried water molecules in crystal structures of acetylcholinesterase from Torpedo California. J Mol Biol 2000;296:713–735.
7. Likic VA, Prendergast FG. Dynamics of internal water in fatty acid binding protein: computer simulations and comparison with experiments. Proteins 2001;43:65–72.
8. Lau EY, Bruice TC. The active site dynamics of 4-chlorbenzoyl-CoA dehalogenase. Proc Natl Acad Sci USA 2001;98:9527–9532.
9. Naidoo KJ, Kuttel M. Water structure about the dimer and hexamer repeat units of amylose from molecular dynamics: computer simulations. J Comput Chem 2001;22:445–456.
10. Belissent–Funel M-C. Hydration in protein dynamics and function. J Mol Liq 2000;84:39–52.
11. Morgan DO. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol 1997;13:261–291.
12. Meijer L, Leclerc S, Leost M. Properties and potential applications of chemical inhibitors of cyclin-dependent kinases. Pharmacol Therap 1999;82:279–284.
13. De Bondt HL, Rosenblatt J, Jancarik J, Jones HD, Morgan Do, Kim SH. Crystal structure of cyclin-dependent kinase 2. Nature 1993;363:595–602.
14. Gray NS, Wodicka L, Thunnissen AMWH, Norman TC, Kwon SJ, Espinoza FH, Morgan DO, Barnes G, Leclerc S, Meijer L, Kim SH, Lockhart DJ, Schultz PG. Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. Science 1998;281: 533–538.
15. Lawrie AM, Noble MEM, Tunnah P, Brown NR, Johnson LN, Endicott JA. Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. Nat Struct Biol 1997;4:796–801.
16. Schulze-Gahmen U, Brandsen J, Jones HD, Morgan D, Meijer L, Veselý J, Kim S-H. Multiple modes of ligand recognition: crystal structures of cyclin-dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopenthenyladenine. Proteins 1995;22:378–391.
17. De Azevedo WF, LeClerc S, Meijer L, Havlíček L, Strnad M, Kim S-H. Inhibition of cyclin-dependent kinases by purine analogues. Eur J Biochem 1997;243:518–526.
18. De Azevedo WF, Mueller-Dieckman H-J, Schultze-Gahmen U, Worland PJ, Sausville EA, Kim SH. Structural basis for specifity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. Proc Natl Acad Sci USA 1996;93:2735–2740.
19. Sielecki TM, Johnson TL, Liu J, Muckelbauer JK, Grafstrom RH, Cox S, Boylan JF, Burton CR, Chen H, Smallwood A, Chang C-H, Boisclair M, Benfield PA, Trainor GL, Seitz SP. Quinazolines as cyclin dependent kinase inhibitors. Bioorg Med Chem Lett 2001;11: 1157–1160.
20. Hoessel R, LeClerc S, Endicott JA, Noble M, Lawrie A, Tunnah P, Leost M, Damiens E, Domonique M, Marko D, Niederberger E, Tang W, Eisenbrand G, Meijer L. Indirubin, the active constituent of a Chinese antileukemia medicine inhibits cyclin-dependent kinases. Nat Cell Biol 1999;1:60–67.
21. Russo AA, Jeffrey PD, Patten AK, Massagué JNPP. Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A0CDK2 complex. Nature 1996;382:325–331.
22. Davies TG, Endicott JA, Noble MEM, Johnson LN, Arris CE, Bentley J, Boyle TF, Calvert AH, Curtin NJ, Jewsbury PJ, Gibson AE, Golding BT, Griffin RJ, Hardcastle I, Mesguiche V, Parsons R, Whitfield H, Newell DR. Structural and thermodynamic validation of inactive CDK2 as a template for structure-based drug design. Cell Mol Biol Lett 2001;6:514–515.
23. Naumann T, Matter H. Structural classification of protein kinases using 3D molecular interaction field analysis of their ligand binding sites: target family landscapes. J Med Chem 2001;45:2366–2378.
24. Case DA, Pearlman DA, Caldwell JW, Cheatham TE III, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crosley M, Radmer RJ, Seibel GL, Singh UC, Weiner PK, Kollman PA. AMBER 6. San Francisco: University of California; 1999.
25. Cornell WD, Cieplak P. Bayly CI, Gould IR, Merz JKM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A 2nd generation force-field for simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
26. Vriend G. WHAT IF, 5.0. Heidelberg: EMBL; 1997.
27. Hooft RWW, Sander C, Vriend G. Positioning hydrogen atoms by optimizing hydrogen bond network in protein structures. Proteins 1996;26:363–376.
28. Otyepka M, Kříž Z, Koča J. Dynamics and binding modes of free CDK2 and its two complexes with inhibitors studied by computer simulations. J Biomol Struct Dynam 2002;20:141–154.
29. Essmann U, Perera L, Berkowitz ML, Darden TA, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103:8577–8593.
30. Bernendsen HC, Postma JPM, van Gunsteren WF, DiNola ARHJ. Molecular dynamics with coupling to an external bath. J Comp Phys 1984;81:3684–3690.
31. Ryckaert JP, Ciccotti G, Berendsen HC. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alcanes. J Comp Phys 1977;23:327–341.
32. Makarov VA, Andrews BK, Smith PE, Pettitt MB. Residence times of water molecules in the hydration sites of myoglobin. Biophys J 2000;79:2966–2974.
33. Ferrin TE, Huang CC, Jarvis LE, Langridge R. The MIDAS display system. J Mol Graph 1988;6:13–27.
34. Rittenhouse R. Unpublished software. 2000.
35. Burgess J. Ions in solution—basic principles of chemical interactions. Chichester, UK: Horwood; 1988.
36. Lynden-Bell RM, Rasaiah JC, Noworyta JP. Using simulation to study solvation in water. Pure Appl Chem 2001;73:1721–1731.
37. Pal SK, Peon J, Zewail AH. Biological water at the protein surface: dynamical solvation probed directly with femtosecond resolution. Proc Natl Acad Sci USA 2002;99:1763–1768.
38. Denisov VP, Halle B. Protein hydration dynamics in aqueous solution: a comparison of bovine pancreatic trypsin inhibitor and ubiquitin by oxygen-17 spin relaxation dispersion. J Mol Biol 1995;245:682–697.
39. Denisov VP, Halle B. Hydrogen exchange and protein hydration: the deuteron spin relaxation dispersions of bovine pancreatic trypsin inhibitor and ubiquitin. J Mol Biol 1995;245:698–709.
40. Desiraju GR, Steiner T. The weak hydrogen bond in structural chemistry and biology. Oxford, UK: Oxford University Press; 1999.
41. Davies TG, Pratt DJ, Endicott JA, Johnson LN, Noble MEM. Structure-based design of cyclin-dependent kinase inhibitors. Pharmacol Therap 2002;93:125–133.
42. Kryštof V, Strnad M. Inhibitors of cyclin-dependent kinases. Chem Listy 2001;95:295–300.