J Mol Model (2005) 11: 509–515
DOI 10.1007/s00894-005-0263-8

ORIGINAL PAPER

Yongjun Jiang · Jianwei Zou · Chunshan Gui

Study of a ligand complexed with Cdk2/Cdk4 by computer simulation

Received: 15 October 2004 / Accepted: 28 February 2005 / Published online: 1 June 2005

Abstract Cyclin-dependent kinases (Cdks) play important roles in the regulation of the cell cycle. Their inhibitors have entered clinical trials to treat cancer. Very recently, Davis et al. (Nat Struct Biol 9:745–749, 2002) have found a ligand NU6102, which has a high affinity with cyclin-dependent kinase 2 (K_i = 6 nM) but a low affinity with cyclin-dependent kinase 4 (K_i = 1,600 nM). To understand the selectivity, we use homology modeling, molecular docking, molecular dynamics and free-energy calculations to analyze the interactions. A rational 3D model of the Cdk4–NU6102 complex is built. Asp86 is a key residue that recognizes NU6102 more effectively with Cdk2 rather than Cdk4. Good binding free energies are obtained. Energetic analysis reveals that van der Waals interaction and nonpolar contributions to solvent are favorable in the formation of complexes and the sulfonamide group of the ligand plays a crucial role for binding selectivity between Cdk2 and Cdk4.

Keywords Homology modeling · Molecular docking · Molecular dynamics · MM-PBSA · Binding selectivity

Introduction

The cyclin-dependent kinases (Cdks) are a family of Ser/Thr kinases that play a central role in the regulation of the cell cycle [1, 2]. They are increasingly being identified as important targets for therapeutic intervention in cancer [3, 4]. A number of small-molecule inhibitors of Cdks have been identified and described in recent reviews [5–7]. Flavopiridol is the most successful flavonoid anticancer agent in its inhibition of multiple Cdks such as Cdk1, Cdk2 and Cdk4 [8]. Since these enzymes are involved at different stages of the cell cycle, more selective inhibitors would therefore provide a broader palette of biological profiles.

Computational techniques in computer-aided drug design are an important tool for illustrating receptor–ligand interactions. Some computational approaches were recently used to study Cdk–ligand recognition [9–13]. Recently, Davies et al. [14] designed a new inhibitor, NU6102, which has a high affinity with Cdk2 (K_i = 6 nM) but a low one for Cdk4 (K_i = 1,600 nM). To understand the origin of the difference, we have used homology modeling, molecular docking, molecular dynamics (MD) and free-energy calculations to analyze the interactions between the inhibitor and Cdks.

Methods

Homology modeling

Although the structure of Cdk4 has not been solved, the level of sequence homology between Cdk4 and Cdk2 is quite high (sequence identity 45.6%) and it seemed reasonable to construct a Cdk4 model based on the Cdk2 structure. Encouragingly, Honma et al. [15, 16] have applied a new de novo design strategy using a Cdk4 homology model according to an activated form of Cdk2 and found some potent Cdk4 inhibitors. In the present paper, the 3D structure of Cdk4 was built using the Homology module of InsightII 2000 [17] based on a monomer of the X-ray structure of the NU6102–Cdk2 complex (pdb code 1H1S) excluding cyclin A. The model was optimized using the molecular mechanics method with the fol-

Y. Jiang (✉) · J. Zou
Key Laboratory for Molecular Design and Nutrition Engineering,
Ningbo Institute of Technology,
Zhejiang University, Ningbo, 315104,
P.R. China
E-mail: yjjiang@nit.net.cn

C. Gui
Center for Drug Discovery and Design,
Shanghai Institute of Materia Medica,
200031, P.R. China

lowing parameters: a distance-dependent dielectric constant of 5.0; nonbonded cutoff 10 Å ; Amber force field [26] and Kollman all-atom charges [17]; and conjugate gradient minimized until the energy gradient RMS−1 . The final minimization structure was used to molecular docking.

Automated molecular docking

The interaction mode between NU6102 and Cdk4 was explored by the docking program AutoDock 3.0 [18]. The Lamarckian genetic algorithm (LGA) [19] was used to deal with the inhibitor–protein interactions. The whole docking process can be described as follows. First, the Cdk molecule was checked for polar hydrogens and assigned partial atomic charges. The PDBQs file was created, and the atomic solvation parameters were also assigned for the macromolecules. Meanwhile, all the torsion angles of the inhibitors to be explored during molecular docking were defined. This allowed a conformational search of inhibitors during the process of docking. Second, a 3D grid was created by the AutoGrid algorithm [19] to evaluate the binding energies between the inhibitor and the protein. Third, a series of docking parameters were set up. The number of generations, energy evaluations and docking runs were set to 50,000, 1,500,000 and 20, respectively. The docking parameters were tested by docking NU6102 into Cdk2. Fourth, the model Cdk4 was aligned on the X-ray structure of the Cdk2–NU6102 complex. Cdk2 was excluded, and the position of NU6102 was used as the initial position for docking. The lowest-energy docked complex was found in the majority of similar docking conformations. Finally, the lowest-energy docked complex was selected for further study.

Molecular dynamics simulation

Force-field parameters

Since comparison of the X-ray structures for variant states of Cdk2 including inactive and active state reveals only very small differences in the vicinity of inhibitor-binding site between different activation states of Cdk2, the monomer of phosphorylate Cdk2 was mutated into the normal Thr at the 160th residue. The force-field parameters for Cdk2 and Cdk4 were obtained from the parm98 force field in Amber7 [20].

The ligand NU6102 was fully minimized by the AM1 method and electrostatic potentials computed at the HF 6-31G* level in the Gaussian 98 program [21]. The RESP fitting technique in AMBER was used to determine the partial charges. The missing van der Waals nonbonded parameters of the ligand were generated by the Antechamber program.

Molecular dynamics simulation

All MD simulations were carried out at 298 K with the ligand bound to the protein and with a cap of waters around the complex filled up to 25 Å from the center of mass of the ligand. The explicit solvent model TIP3P was used for water. SHAKE was applied to all bonds involving hydrogen and a time step of 2 fs was used [22]. The nonbonded pairs were updated every 30 steps with a cutoff of 12 Å for the nonbonded interactions. All residues within 15 Å and all waters were allowed to move, and the rest was fixed. Prior to the MD simulations, 2,000 steps of minimization with the steepest descent method were used to remove high-energy interatomic contacts. Each system was run for 600 ps and 50 snapshots were collected every 4 ps during the final 200 ps for each of the MD trajectories for the free-energy analysis.

Free energy calculation

The free energy was approximated as the sum of the molecular mechanical energy in vacuo, the solvation free energy [23, 24]:

$$G = E_{MM} + G_{solv} - TS$$

where

$$G_{solv} = G_{polv} + G_{np}.$$

The total molecular mechanical energies, E_{MM}, include bond, angle, dihedral, E(BADH), van der Waals, and electrostatics terms. The electrostatic component of the solvation free energy, G_{pol}, was computed by the finite-difference Poisson–Boltzmann approximation [25]. In this approach, the solvent was treated as a continuum model of high dielectric ($\varepsilon = 80$) and the solute as a low-dielectric medium ($\varepsilon = 1$) with embedded charges taken from the Cornell et al. [26] force field. The radii of the atoms were taken from the PARSE parameter set [27]. The salt effect was neglected. We defined a 0.5 Å gird spacing with lattice dimensions extending 20% beyond the solute dimensions and required 1,000 linear iterations. The hydrophobic contribution to the solvation free energy was determined with a solvent-accessible surface-area-dependent term [27]. The solvent-accessible surfaces (SAS) were calculated with Paul Beroza's molsurf program, which is based on analytical ideas by Mike Connolley [28]. G_{np} was calculated as $0.00542 \times SAS + 0.92$ [27].

Normal-mode analysis

We estimated the conformational entropy contributions (translation, rotation and vibration) to the binding free energy using normal-mode analysis. As they are very time-consuming, to simplify the calculations, the residues further than 10 Å from the ligand in the two ter-

minals of the Cdks from the final MD snapshots were excluded and the systems had about 3,500 atoms and the complexes did not have unbonded valences. The complexs, receptors and ligands were minimized using a conjugate gradient strategy for 50,000 steps in the presence of a distance-dependent dielectric (4*rij*) using the sander program in AMBER. The entropy for each structure was calculated using classical statistical formulae and normal-mode analysis.

Computational mutagenesis

The approach mutated several snapshots from the wild-type trajectory by altering their coordinates [29]. This method was applied to obliterating changes or amino acid deletions such as alanine and glycine mutations. Meanwhile, the mutated systems were not minimized, and it is assumed that no local rearrangements occur with the mutation.

Results and discussion

The sequence-alignment result from the segment pair overlap algorithm is shown in Fig. 1 [30]. The monomer of Cdk2 from NU6102 complexed to phospho-Cdk2-cyclin A was used as a template for modeling the active Cdk4 structure. The complex structure of Cdk4 with NU6102 by molecular docking is shown in Fig. 2. Figure 3a, b shows the interacting modes of compound NU6102 with Cdk2 and Cdk4, respectively, derived by the LIGPLOT program [31]. From Fig. 3b, ten residues form a hydrophobic pocket. Following a similar binding pattern, hydrogen bonding is one important characteristic of the interaction between NU6102 and Cdk2 or Cdk4. Hydrogen bonds are formed between the purine core and the backbone of residues Leu83 and Glu81 in the Cdk2–NU6102 complex. Instead, the ligand donates two hydrogen bonds to Glu94 (rNH–O = 2.85 Å) and Val96 (rNH2–O = 2.59 Å) in the Cdk4 complex, which is consistent with the results Honma et al. [15, 16], who predicted structural requirements for Cdk4 model. Furthermore, because the NH2 group of the sulfonamide donates a hydrogen bond to a side chain oxygen of Asp 86 (RNH2–O = 2.9 Å), and one sulfonamide oxygen accepts a hydrogen bond from the backbone nitrogen of Asp 86 (rNH–O = 3.1 Å), Cdk2 has a high affinity for NU6102. However, in the Cdk4–NU6102 complex, from our model, the sulfonamide group is exposed outside the binding pocket of Cdk4.

MD simulations of the NU6102 complex with Cdk2 or Cdk4 were performed with a spherical cap of explicit waters and a "belly", in which atoms are allowed to move during the simulations. The root-mean-squared distances (RMSDs) of the ligand from the initial structures are reported in Table 1. The average RMSD fitted to the starting structure of the ligand in the Cdk2 complex is 1.112 Å . However, the RMSD fitted to the starting structure of the ligand in the Cdk4 complex is 2.149 Å . The NU6102 in the Cdk4 complex is more flexible than that in the Cdk2 complex, which agrees with the structural information of these two complexes.

The results from energetic analyses of 50 equally spaced snapshots taken from each of the two MD simulations are summarized in Table 2. The calculated free energies of the Cdk4–NU6102 complex ($\Delta G_{\text{binding}} = -8.67$ kcal mol^{-1}) and of the Cdk2–NU6102 complex formation ($\Delta G_{\text{binding}} = -10.65$ kcal mol^{-1}) agree fairly well with those from experiment (−7.9 and −11.29 kcal mol^{-1}, respectively). The favorable formation of these complexes is driven by van der Waals contributions (ΔE_{vdw} values) and nonpolar contributions to solvation (ΔG_{np}). As demonstrated by numerous studies [22–28], electrostatics generally disfavor the docking of ligand and receptor molecules because the unfavorable change in the electrostatics of solvation is mostly compensated by the favorable electrostatics

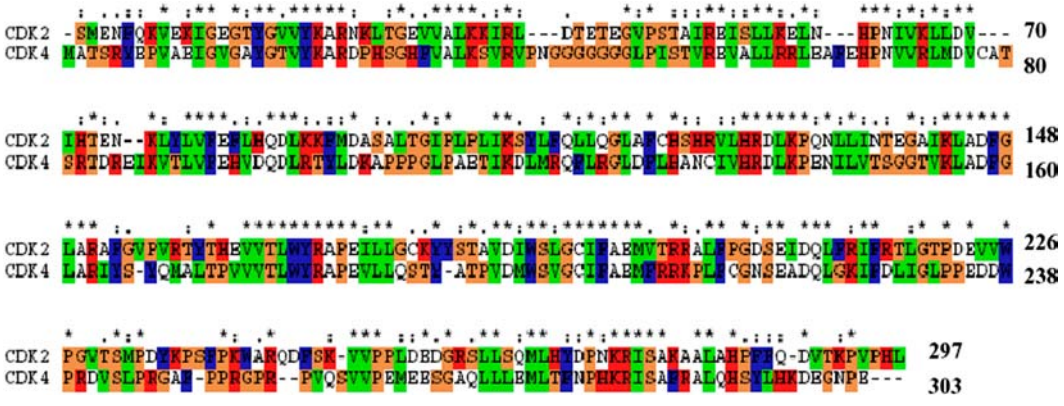


Fig. 1 The sequence alignments of Cdk2 and Cdk4 generated by CLUSTALW. In the sequences, an *asterisk* (*) indicates an identical or conserved residue; a *colon* (:) indicates conserved substitutions; a *stop* (.) indicates semi-conserved substitutions

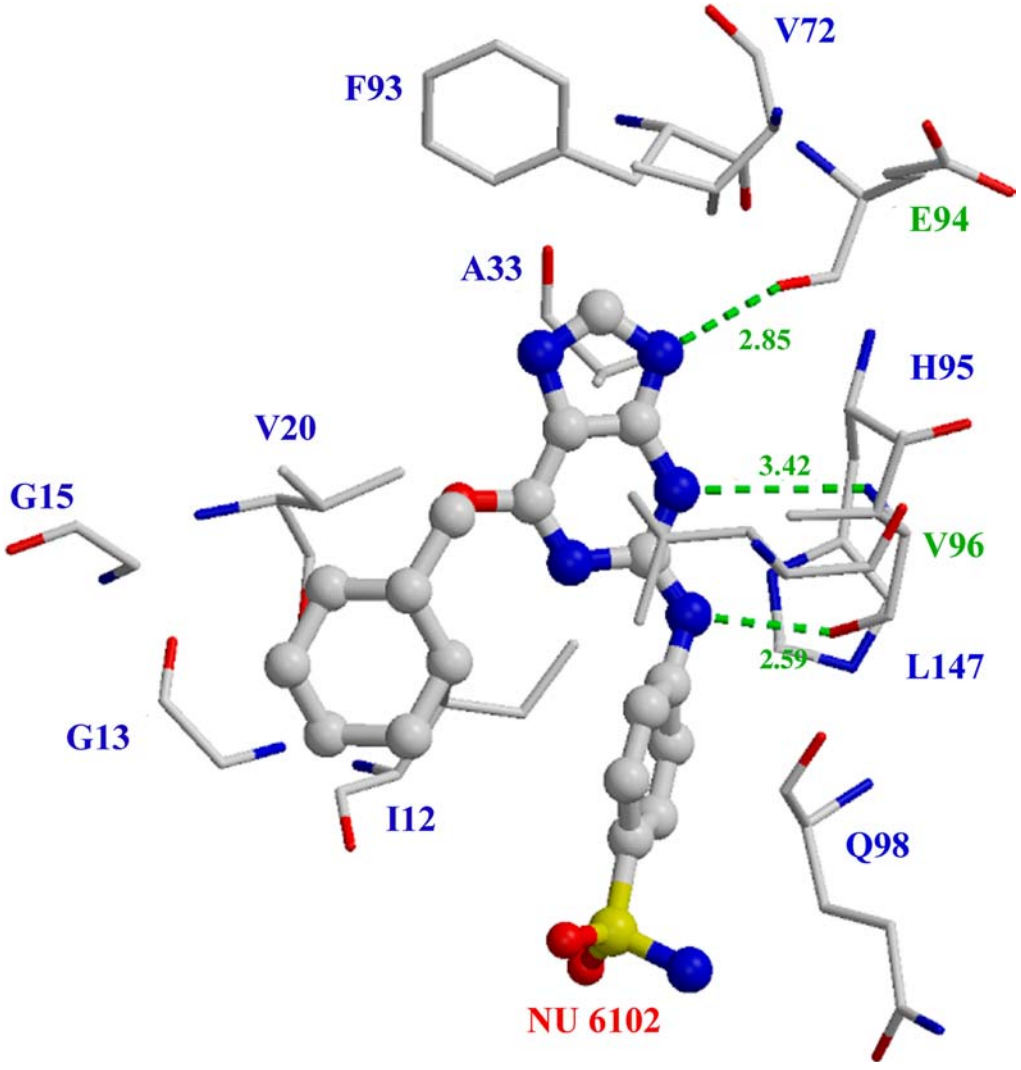


Fig. 2 Predicted binding model of Cdk4–NU6102 by homology modeling and molecular docking, hydrogen bonds are drawn as *dotted lines*

within the resulting ligand–receptor complex. Here, the total electrostatic energy contributions (ΔG_{pbele}) for both Cdk2 and Cdk4–NU6102 complex formation are unfavorable, with values of 26.27 and 31.48 kcal mol^{-1}, respectively. Since the sulfonamide forms another two hydrogen bonds with Asp86 in the Cdk2–NU6102 complex, the electrostatic interaction in Cdk2–NU6102 ($\Delta E_{\text{ele}} = -49.21$ kcal mol^{-1}) is more favorable than in Cdk4–NU6102 ($\Delta E_{\text{ele}} = -30.87$ kcal mol^{-1}).

As shown above, the sulfonamide group plays an important role in recognition by Cdk2 and Cdk4. Our energetic analyses of individual group's interactions clearly elucidate the structural character (Table 3). The contribution of the purine core is most favorable for ligand binding, as excepted (−42.89 kcal mol^{-1} for Cdk2–NU6102, −43.53 kcal mol^{-1} for Cdk4–NU6102, respectively). The sum E_{mm} of the sulfonamide group is favorable for Cdk2–NU6102 ($E_{\text{mm}} = -11.28$ kcal mol^{-1}) but unfavorable for Cdk4–NU6102 ($E_{\text{mm}} = 14.89$ kcal mol^{-1}), which also indicates that the sulfonamide group is thus a major binding-selectivity determinant.

To investigate the role of some residues in the ligand–protein interaction, we applied a computational mutagenesis replacing these residues by alanines. Results from our energetics of the mutant complexes are shown in Tables 4 and 5. In the Cdk2–NU6102 complex, mutation of ILE11 to Ala was unfavorable by 2.42 kcal mol^{-1} . The mutation induced a reduction of 2.85 kcal mol^{-1} in E_{vdw}. Like ILE11, Leu135 is also unfavorable by 2.80 kcal mol^{-1} replacement by Ala resulting in a reduction of 2.85 kcal mol^{-1} in E_{vdw}. However, E81 and D86 form hydrogen bonds with NU6102, so their mutations are found to be unfavorable for E_{ele}, as shown by $\Delta\,\Delta E_{\text{ele}} = 5.7$ kcal mol^{-1} for E81/Ala and $\Delta\,\Delta E_{\text{ele}} = 18.64$ kcal mol^{-1} for D86/Ala.

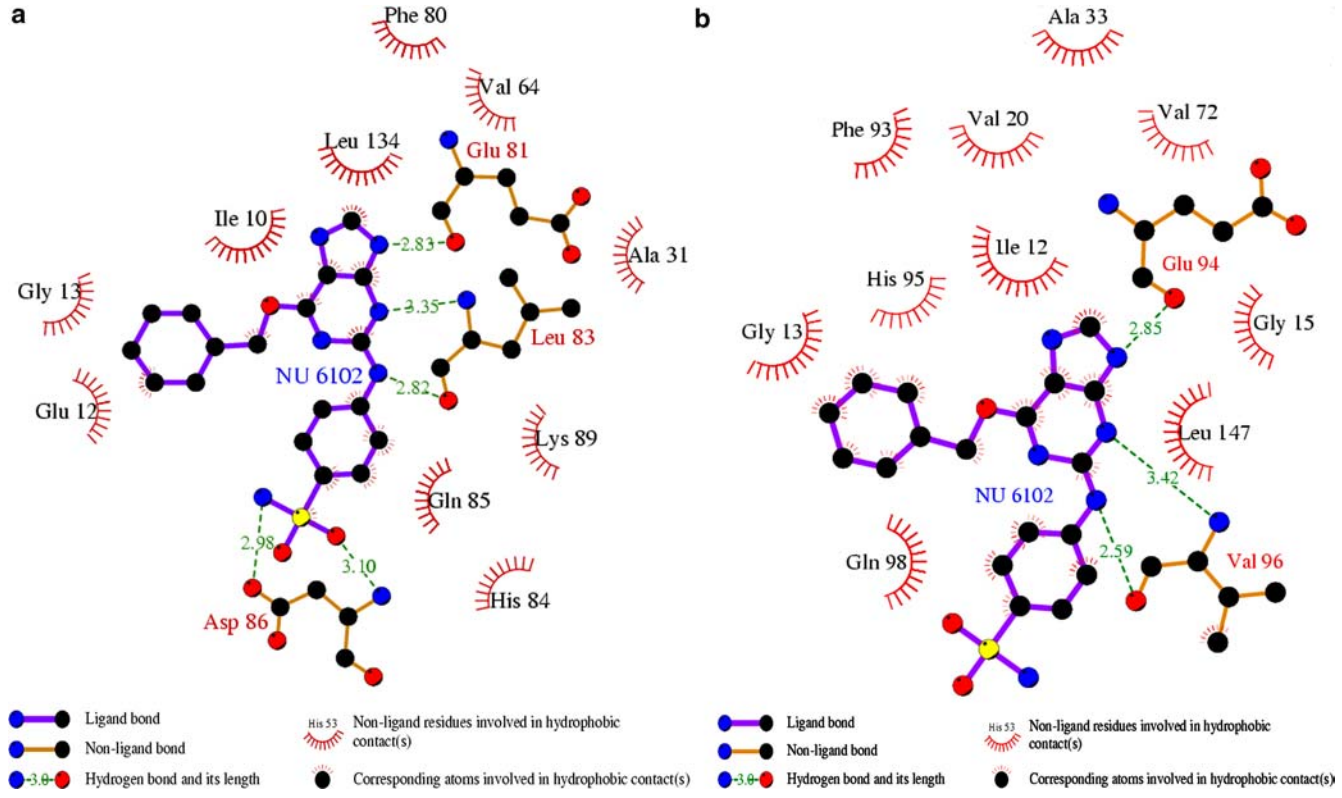


Fig. 3 **a** Two-dimensional representative for the interacting model of NU6102 complexed with the Cdk2 from an X-ray structure by LIGPLOT. **b** Two-dimensional representative for the interacting model of NU6102 complexed with the Cdk4 from a predicted structure by LIGPLOT

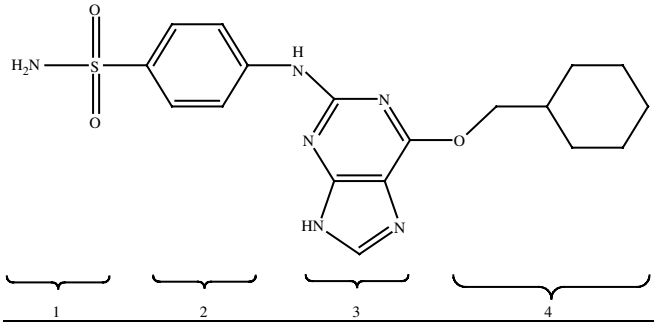
Table 1 RMS deviation (Å) of the ligand in simulations from its starting structure

Complex	Average	Maximum	Minimum
Cdk2	1.112	1.699	0.225
Cdk4	2.149	2.549	0.713

Table 2 Binding free energies (kcal mol^{-1}) of Cdk4, Cdk2 and their ligand NU6102

Contribution	Cdk4–NU6102		Cdk2–NU6102	
	Mean	STD	Mean	STD
ΔE_{ele}	−30.87	3.94	−49.21	7.24
ΔE_{vdw}	−44.76	3.30	−47.20	3.99
ΔE_{gas}	−75.63	4.42	−96.42	7.22
ΔG_{np}	−2.87	0.17	−3.10	0.14
ΔG_{pb}	62.35	3.56	75.49	7.30
ΔG_{pbele}	31.48	4.01	26.27	3.89
$\Delta\ \Delta G_{pb,\ tot}$	−16.15	4.01	−24.03	3.68
TS	7.48	13.38		
$\Delta G_{binding}$	−8.67	−10.65		
Experiment*	−7.90	−11.29		

Table 3 Molecular mechanic interaction energy (kcal mol^{-1}) between Cdks and NU6102



Group	Cdk2–NU6102 (kcal mol^{-1})	Cdk4–NU6102 (kcal mol^{-1})
$E_{mm}1$	−11.28	14.89
$E_{mm}2$	−13.76	−11.84
$E_{mm}3$	−42.89	−43.53
$E_{mm}4$	−12.78	−14.66

Energetic analyses are computed by Anal program [19, 32].

Although Asp is conserved through the whole Cdk family, our calculations indicate that D86 in Cdk2 is an important residue for recognizing the inhibitor. However, the corresponding residue of D99 in Cdk4 does not act on the small molecule. In the Cdk4–NU6102 complex, since the hydrogen bond between NU6102 and Glu94 is broken upon mutation of Glu to Ala, E_{ele}

Table 4 Relative binding free energies (kcal mol^{-1}) between wild and mutant Cdk2–NU6102 complex

Contribution	I11A	E81A	L83A	D86A	L134A
$\Delta\ \Delta E_{ele}$	−0.13	5.78	−0.15	18.64	−0.03
$\Delta\ \Delta E_{vdw}$	2.85	0.22	0.82	−1.52	3.73
$\Delta\ \Delta E_{gas}$	2.73	6.02	0.67	17.13	3.69
$\Delta\ \Delta G_{np}$	0.06	0.05	0.00	0.06	0.08
$\Delta\ \Delta G_{pb}$	−0.46	−4.66	−0.24	−14.38	−1.09
$\Delta\ \Delta G_{pb,\ tot}$	2.42	1.45	0.43	2.80	2.80

Table 5 Relative binding free energies (kcal mol^{-1}) between wild and mutant Cdk4–NU6102 complex

Contribution	E94A	V96A	D99A	I12A	V20A
$\Delta\ \Delta E_{ele}$	6.25	0.13	1.28	0.09	0.15
$\Delta\ \Delta E_{vdw}$	0.24	1.15	0.95	2.80	1.35
$\Delta\ \Delta E_{gas}$	6.48	1.28	2.23	2.89	1.49
$\Delta\ \Delta G_{np}$	0.00	0.00	0.07	0.25	0.06
$\Delta\ \Delta G_{pb}$	−5.18	−0.34	−1.95	−1.73	−0.24
$\Delta\ \Delta G_{pb,tot}$	1.30	0.95	0.35	1.41	1.32

is unfavorable by 6.25 kcal mol^{-1} and the total binding energy is unfavorable by 1.30 kcal mol^{-1}. Surprisingly, the binding free energies are not changed as expected with the mutations of L83 in Cdk2–NU6102 and V96 in Cdk4–NU6102.

Conclusion

Understanding protein–ligand interactions is essential for designing more selective and potent inhibitors. In this paper, we have presented an approach that combines homology modeling, molecular docking, molecular dynamics and Poisson–Boltzmann free-energy simulations to understand the binding selectivity of NU6102 complexed with Cdk2 and Cdk4. A reasonable 3D structure of Cdk4 was constructed based on an activated monomer of Cdk2. By molecular docking, a complex model provides a satisfactory explanation for the binding mechanism of Cdk4 with NU6102. Our study suggests the origin of binding selectivity of NU6102 complexed with Cdk2/Cdk4, which is mainly the sulfonamide group exposed outside binding pocket of Cdk4 but forming a hydrogen bond with Asp 86 in Cdk2–NU6102. Good binding free energies are obtained by MM/PBSA, and the calculations suggest that van der Waals interactions and nonpolar contributions to solvation provide the driving force for binding. Ala scanning reveals some key residues involved in the binding of complex.

Acknowledgments The authors thank Professor Arthur J. Olson for his kindness in offering us the AutoDock 3.0.3 program. We also thank Dr. Holger Gohlke for his help with free energy calculations. Jiang YJ acknowledges Prof. Jiang Hualiang.

References

1. Hunter T, Pines J (1994) Cell 79:573–582
2. Sherr C (1996) Science 274:1672–1677
3. Garrett MD, Fattaey A (1999) Curr Opin Genet Dev 9:104–111
4. Webster KR (1998) Expert Opin Invest Drugs 7:865–887
5. Sielecki TM, Boylan JF, Beneld PA, Trainor GL (2000) J Med Chem 43:1–18
6. Toogood PL (2001) Med Res Rev 21:487–498
7. Huwe A, Mazitschek R, Giannis A (2003) Angew Chem Int Ed 42:2122–2138
8. Webster KR (2000) Chem Res Toxicol 13:940–943
9. Sims PA, Wong CF, McCammon JA (2003) J Med Chem 46:3314–3325
10. Otyepka M, Kriz Z, Koca J (2002) J Biomol Struct Dyn 20:141–154
11. Kriz Z, Otyepka M, Bartova I, Koca J (2004) Proteins 55:258–274
12. Wu G, Vieth M (2004) J Med Chem 47:3142–3148
13. McInnes C, Wang S, Anderson S, O'Boyle J, Jackson W, Kontopidis G, Meades C, Mezna M, Thomas M, Wood G, Lane DP, Fischer PM (2004) Chem Biol 11:525–534
14. Davies TG, Bentley J, Arris CE, Boyle FT, Curtin NJ, Endicott JA, Gibson AE, Golding BT, Griffin RJ, Hardcastle IR, Jewsbury P, Johnson LN, Mesguiche V, Newell DR, Noble ME, Tucker JA, Wang L, Whitfield HJ (2002) Nat Struct Biol 9:745–749
15. Honma T, Hayashi K, Aoyama T, Hashimoto N, Machida T, Fukasawa K, Iwama T, Ikeura C, Ikuta M, Suzuki-Takahashi I, Iwasawa Y, Hayama T, Nishimura S, Morishima H (2001) J Med Chem 44:4615–4627
16. Honma T, Yoshizumi T, Hashimoto N, Hayashi K, Kawanishi N, Fukasawa K, Takaki T, Ikeura C, Ikuta M, Suzuki-Takahashi I, Hayama T, Nishimura S, Morishima H (2001) J Med Chem 44:4628–4640
17. Insight II, version 2000 (2002) Molecular Simulations Inc., San Diego, CA
18. Morris GM, Goodsell DS, Huey R, Hart WE, Halliday S, Belew R, Olson AJ (1999) Autodock Version 3.0.3. Molecular Graphics Laboratory, Department of Molecular Biology, The Scripps Research Institute
19. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
20. Case DA, Pearlman DA, Caldwell JW, Cheatham TE, Wang J, Ross WS, Simmerling C, Darden T, Merz KM, Stanton RV, Cheng A, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer R, Duan Y, Pitera J, Seibel GL, Singh UC, Weiner P, Kollman P (2002) AMBER: Version 7. University of California, San Francisco
21. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratman RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu C, Liashenko A, Piskorz P, Komaromi, I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Gonzales C, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98, revision A.7. Gaussian Inc., Pittsburg, PA
22. Ryckaert JP, Ciccotti G, Berendsen HJC (1997) J Comp Phys 23:327–341
23. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998) J Am Chem Soc 120:10629–10633
24. Smith KC, Honig B (1994) Proteins 18:119–132
25. Sharp KA, Honig B (1990) Annu Rev Biophys Biophys Chem 19:301–332

26. Cornell WD, Cieplak P, Bayly CI, Could IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
27. Sitkoff D, Sharp KA, Honig B (1994) J Phys Chem 98:1978–1988
28. Connolly ML (1983) J Appl Cryst 16:548–558
29. Massova I, Kollman PA (1999) J Am Chem Soc 121:8133–8143
30. Higgins DG, Sharp PM (1988) Gene 73:237–244
31. Wallace AC, Laskowski RA, Thornton JM (1995) Protein Eng 8:127–134
32. Chong LT, Duan Y, Wang L, Massova I, Kollman PA (1999) Proc Natl Acad Sci USA 96:14330–14335