

Understanding and modulating cyclin-dependent kinase inhibitor specificity: molecular modeling and biochemical evaluation of pyrazolopyrimidinones as CDK2/cyclin A and CDK4/cyclin D1 inhibitors

Karen A. Rossi^{a,*}, Jay A. Markwalder^a, Steven P. Seitz^a, Chong-Hwan Chang^a, Sarah Cox^a, Michael D. Boisclair^b, Leonardo Brizuela^c, Stephen L. Brenner^a & Pieter F. W. Stouten^d

^aPharmaceutical Research Institute, Bristol-Myers Squibb Company, 5400 Princeton NJ, 08543, USA; ^bOSI Pharmaceuticals, 1 Bioscience Park Drive, Farmingdale, NY, 11735, USA; ^cDepartment of BCMP, Harvard Medical School, 240 Longwood Ave., SGM. Room 211, Boston, MA, 02115, USA; ^dNerviano Medical Sciences, Viale Pasteur 10, 20014, Nerviano (MI), Italy

Received 8 January 2004; accepted in revised form 28 January 2005
© Springer 2005

Key words: ATP-competitive, cancer, cyclin-dependent kinase (CDK), docking, homology, selectivity

Summary

Cyclin-dependent kinases (CDKs) play a key role in regulating the cell cycle. The cyclins, their activating agents, and endogenous CDK inhibitors are frequently mutated in human cancers, making CDKs interesting targets for cancer chemotherapy. Our aim is the discovery of selective CDK4/cyclin D1 inhibitors. An ATP-competitive pyrazolopyrimidinone CDK inhibitor was identified by HTS and docked into a CDK4 homology model. The resulting binding model was consistent with available SAR and was validated by a subsequent CDK2/inhibitor crystal structure. An iterative cycle of chemistry and modeling led to a 70-fold improvement in potency. Small substituent changes resulted in large CDK4/CDK2 selectivity changes. The modeling revealed that selectivity is largely due to hydrogen-bonded interactions with only two kinase residues. This demonstrates that small differences between enzymes can efficiently be exploited in the design of selective inhibitors.

Abbreviations: (K2A) – CDK2/cyclin A complex; (K2E) – CDK2/cyclin E complex; (K4D1) – CDK4/cyclin D1 complex.

Introduction

One of the most important and fundamental processes in biology is the division of cells mediated by the cell cycle. This process ensures the controlled production of new generations of cells with defined biological function. The cell cycle can be divided into four distinct phases [1–4]: the first gap phase (G1) prepares the cell for DNA replication. In the

synthesis phase (S) a second copy of the entire genome is synthesized. During the second gap phase (G2) the cell prepares for division, and finally in the mitosis phase (M) the two copies of the DNA separate and two genetically identical daughter cells result. This cell division process is a highly regulated phenomenon and responds to a diverse set of cellular signals both within the cell and from external sources. A complex network of tumor promoting and suppressing gene products are key components of this cellular signaling process. Over-expression of the tumor promoting components or

*To whom correspondence should be addressed. Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 5400, Princeton, NJ 08543-4000.; E-mail: karen.rossi@bms.com

the loss of the tumor suppressing products will lead to unregulated cellular proliferation and the generation of tumors [5, 6].

Cyclin dependent kinases (CDKs) play a key role in regulating the cell cycle machinery. These complexes consist of two components: a catalytic subunit (the kinase) and a regulatory subunit (the cyclin). To date, several kinase subunits (CDKs 1 through 11) have been identified along with several regulatory subunits (cyclins A through L). Each kinase associates with a specific regulatory partner and together make up the active catalytic moiety. Each transition of the cell cycle is regulated by particular CDK complexes: G1/S by CDK2/cyclin E (K2E), CDK4/cyclin D1 (K4D1) and CDK6/cyclin D2; S/G2 by CDK2/cyclin A (K2A) and CDK1/cyclin A; and G2/M by CDK1/cyclin B. The coordinated activity of these kinases guides the cell through the replication process and ensures the vitality of each subsequent generation [7, 8].

An increasing body of evidence has shown a link between tumor development and CDK-related malfunctions. Overexpression of the cyclin regulatory proteins and subsequent kinase hyperactivity have been linked to several types of cancers [9–11]. More recently, endogenous, highly specific protein inhibitors of CDKs were found to have a major effect on cellular proliferation [12, 13]. These inhibitors include p16^{INK4} (a K4D1 inhibitor), p21^{CIP1} (a general CDK inhibitor), and p27^{KIP1} (a CDK2 inhibitor). A crystal structure of p27 bound to K2A revealed how these proteins effectively inhibit the kinase activity through multiple interactions with the CDK complex [14]. These proteins help regulate the cell cycle through specific interactions with their corresponding CDK complexes. Cells deficient in these inhibitors are prone to unregulated growth and tumor formation. This body of evidence has led to an intense search for small molecule inhibitors of the CDK family as an approach to cancer chemotherapy. Since many cancers are associated with G1 and the G1/S transition [3, 4], one of the primary goals of our cell cycle program was the identification and optimization of chemical entities that inhibit K4D1 specifically.

The focus of this paper is the molecular modeling of pyrazolopyrimidinone inhibitors into the CDK4 ATP-binding pocket, the validation of the binding model using the crystal structure of a CDK2/inhibitor complex, and the use of the

binding models to design more potent enzyme inhibitors and to control K4D1/K2E selectivity. The *in vitro* assays used to identify K4D1 and K2E inhibitors, the identification of the initial lead, and the comparative modeling of CDK4 using CDK2 as a template are also described.

Materials and methods

In vitro kinase reactions

K4D1 and K2E complexes were expressed in insect cells following dual infection by baculovirus, each containing one of the components. Extracts containing K4D1 were prepared as described previously [15]. K2E was purified by chromatography on NTA agarose by virtue of a polyhistidine tag engineered onto the N-terminus of cyclin E, followed by gel filtration chromatography (X. Xu, personal communication). Kinase activity was measured in polypropylene microtiter plates as described previously [16]. Each reaction contained 50 mM Tris (pH 7.6), 10 mM MgCl₂, 1 mM DTT, 10% v/v DMSO, 50 μ M ATP, and 20 μ Ci/mL of [γ -³²P]ATP. K4D1 assays contained 6 μ g GST-Rb per 50 μ L reaction for IC₅₀ determinations, or 20 μ g GST-Rb per 100 μ L reaction for compound screening assays. K2E assays contained 12 μ g GST-Rb per 50 μ L reaction. Kinase reactions were initiated by the addition of 4 units K4D1 or 3 units K2E and allowed to proceed for 15 min (1 unit of enzyme transfers 1 pmol of phosphate to Rb/min). Reactions were stopped by addition of an equal volume of phosphate buffered saline containing 10 mM ATP, 100 mM EDTA, 200 μ g/mL bovine serum albumin, and 0.2% nonidet P40. Phosphate incorporation on Rb was quantitated. Unless otherwise indicated all *in vitro* assays have n of at least 2. Standards for the assays were: CDK4/Cyclin D1, (+/-)-flavopiridol, mean IC₅₀ = 0.089 μ M, std. deviation = 0.019 μ M (*n* = 52); CDK2/Cyclin E, (+/-)-flavopiridol, mean IC₅₀ = 1.05 μ M, std. deviation = 0.31 μ M (*n* = 10); CDK1/Cyclin B, staurosporine, mean IC₅₀ = 0.032 μ M, std. deviation = 0.008 μ M (*n* = 10); HCT116, (+/-)-flavopiridol, mean IC₅₀ = 0.087 μ M, std. deviation = 0.008 μ M (*n* = 10). IC₅₀ values were reproducible to +/-30%. To rapidly identify K4D1 inhibitors, 146,000 compounds from the DuPont corporate database and 16,000 externally acquired

compounds were screened against K4D1 in high-throughput mode. The IC₅₀ values of confirmed active K4D1 inhibitors were subsequently determined in the K4D1 and K2E IC₅₀ assays.

CDK4 homology model building and comparison with CDK2

Since no 3D-structural information is available for CDK4, a homology model was built to dock inhibitors into. The closest CDK4 homolog for which a crystallographic structure exists is CDK2 (46% sequence identity: 134 out of 290 aligned residues are identical). The sequence alignment of CDK4 and CDK2 is shown in Figure 1. CDK2 is not only highly homologous, but over 20 inhibitors have been co-crystallized with CDK2. The details of these crystal structures have been described in great detail elsewhere [17–19]. CDK2 from the K2A complex (PDB entry 1FIN [20, 21]) was used as a template to model the cyclin-bound form of CDK4. The *Modeler* program [22, 23] (Accelrys, San Diego, CA) was used to generate 25 models of CDK4. The Quality Control tool [24] of the *Whatif* program [25] was used to evaluate the models and to select the best one. The ATP binding pocket is defined as the set of residues that have at least one (heavy or hydrogen) atom within a 5 Å radius from any

(heavy or hydrogen) atom of **6** or of the ATP adenosine moiety. For the purpose of the present study, Glu 8 and Lys 9 in CDK2 (Ala 10 and Glu 11 in CDK4) were also regarded as part of the ATP-binding pocket because they are within 7 Å of the catechol moiety of **6** (see Table 2 for structures) and differ significantly between the two enzymes. The most important differences between the CDK2 and CDK4 ATP-binding pockets are detailed in Table 1 and Figure 2. The CDK4 ATP pocket is much more acidic than the CDK2 ATP pocket (by 3 unit charges). Because we have a model structure of CDK4 only, the localization of the main chains, side chains and charges in the CDK4 ATP pocket is not precisely known.

Docking

One or more conformations of a ligand were manually docked in one or more orientations into the CDK4 ATP binding pocket using *InsightII*. Subsequent energy minimizations with the *Discover* program and the *CFF96* force field (Accelrys, San Diego, CA) were carried out for each binding model, keeping the protein rigid but allowing the ligand to be completely flexible. Where appropriate, additional binding modes were generated with the *BForce* flexible

1	11	21	31	41	51	61	71	80
CDK4	MATSRYPVA	EIGVGAYGTV	YKARDPHSGH	FVALKSVRVP	NGGGGGGLP	ISTVREVALL	RRLEAFEHPN	VVRLMDVCAT
	X X X			X				
CDK2	--MENFQKVE	KIGEGTYGVV	YKARNKLTGE	VVALKKIRLD	T--ETEGVP	STAIRESISLL	KEL--NHPN	IVKLLDVIHT
	1	9	19	29	39	46	56	63
81	91	101	111	121	131	141	151	160
CDK4	SRTDREIKVT	LVFEHVDODL	RTYLDKAPPP	GLPAETIKDL	MRQFLRGLDF	LHANCIVHRD	LKPENILVTS	GGTVKLADFG
		X X	XX				X	
CDK2	-----ENKLY	LVFEFLHODL	KKFMDASALT	GIPLPLIKSY	LFQLLQGLAF	CHSHRVLHRD	LKPONLLINT	EGAIKLADFG
	72	78	88	98	108	118	128	138 147
161	170	180	189	199	209	219	229	238
CDK4	LARIYSYQM-	ALTPVVVTLW	YRAPEVLLQS	-TYATPVDMW	SVGCIFAEMF	RRKPLFCGNS	EADQLGKIFD	LIGLPPEDDW
CDK2	LARAFGVFVR	TYTHEVVTW	YRAPEILLGC	KYSTAVDIW	SLGCIFAEMV	TRRALFPGDS	EIDQLFRIFR	TLGTPDEVVW
	148	158	168	178	188	198	208	218 227
239	247	257	267	277	287	297	303	
CDK4	PRDVSLP--R	GAFPPRGRPR	VQSVVPEMEE	SGAQLLLEML	TFNPHKRISA	FRALQHSYLH	KDEGNPE	
CDK2	PGVTSMPPDYK	PSFPKWARQD	FSKVVPPLDE	DGRSLLSQML	HYDPNKRISA	KAALAHPPFQ	DVTKPVP	
	228	238	248	258	268	278	288 294	

Figure 1. CDK4/CDK2 sequence alignment. Identical residues are indicated by vertical bars between the sequences. The nine key residues that differ between CDK2 and CDK4 (see also Table 1) are marked by an X between the sequences. Gaps are indicated by hyphens. The 26 residues that form the ATP binding pocket (i.e., that have at least one atom within 5 Å of any atom of adenosine or **6**) are doubly underlined.

Table 1. Key formal charge differences between CDK2 and CDK4 in the ATP-binding pocket. Histidine can be protonated, but has been assigned a zero charge since it is almost completely buried. The charges of other positive and negative residues in the ATP binding pocket (the same in CDK2 and CDK4) are included in the total absolute charge of the ATP-binding pocket of each enzyme.

CDK2		CDK4	
Residue	Charge	Residue	Charge
Glu 8	-1	Ala 10	0
Lys 9	+1	Glu 11	-1
Glu 12	-1	Val 14	0
Val 29	0	Phe 31	0
Phe 82	0	His 95	0
His 84	0	Asp 97	-1
Lys 88	+1	Arg 101	+1
Lys 89	+1	Thr 102	0
Gln 131	0	Glu 144	-1
Subtotal	+1	Subtotal	-2
Other positive	+2	Other positive	+2
Other negative	-3	Other negative	-3
Total	0	Total	-3

grid-docking method [26] as implemented in the *Affinity* program (Accelrys, San Diego, CA), using SFNS continuum solvation [27]. The *Affinity* program employs Monte Carlo techniques to dock a partially rigid ligand (allowing rotations of up to four bonds) into a rigid binding site. Subsequently, energy minimizations (and optionally Molecular Dynamics simulations) were carried out for each accepted Monte Carlo configuration. In this procedure the ligand is fully flexible as is a shell of protein atoms approximately 6 Å from any heavy atom in the adenosine moiety of ATP. A 3.5 Å buffer region of the protein is flexible, yet restrained to its crystallographic (or model) positions, while the potential due to the remainder of the protein is expressed on a grid [26]. The energy minimizations are run until the RMS force is smaller than 0.001 kcal/mol/Å. The use of *Affinity* has three purposes: (1) it reduces the bias towards preconceived binding modes introduced by the manual docking process; (2) due to the flexibility of the binding site and ligand, more and more favorable docking modes can be found; and (3) due to the use of continuum solvation, the overestimation of charged and polar interactions are reduced to some extent.

Results and discussion

Identification and tentative binding mode determination of the initial lead compound

The initial hits were docked into a CDK4 homology model that was based on the crystal structure of CDK2, which is highly homologous. Similarities and key ATP-binding pocket differences between the two kinases are shown in Figure 1 (sequence alignment), Table 1 (formal charge differences) and Figure 2 (structural differences). Work has been done elsewhere that describes these differences and shows selectivity can be achieved via interacting with different residues of CDK4 [17, 18].

Figure 3 shows three compounds that form the basis of the work described in this paper. Compound **1** is one of the most active hits from the K4D1 high-throughput assay. The similarity of **2** and **1** was realized shortly after the identification of **1** as a K4D1 inhibitor. In order to eliminate the methylsulfide group, a potential metabolic liability present in both **1** and **2**, compound **3** was synthesized [28]. This compound is considered the parent compound of the pyrazolopyrimidinone lead series and was optimized through an iterative cycle of modeling and chemistry. Compound **3** was manually docked into the CDK4 ATP binding pocket and energy minimized. Based on the crystallographic binding mode of ATP to K2A [21] (see Figure 2), the assumption was that the amide moiety of **3** interacts with the CDK4 hinge region, i.e., with the backbone carbonyl of Glu 94 (equivalent to Glu 81 in CDK2) and/or with the backbone amide of Val 96 (CDK2:Leu 83). In that case, two distinct binding modes are possible (Figure 4). Binding mode A (cyan in Figure 4) features two hydrogen bonds between the amide of **1** and Val 96-NH and Glu 94-C=O of the enzyme. In binding mode B (pink in Figure 4) two hydrogen bonds exist between the amide of **3** and the backbone amide of Val 96. After energy minimization of flexible ligands in a rigid active site, the calculated energy of binding mode A is 14 kcal/mol higher than that of binding mode B. The interaction of 5-NH of **3** with Glu 94-C=O in binding mode A, however, is less solvent-exposed than the corresponding interaction with Val 96-C=O in binding mode B. Since buried polar interactions are stronger (due to the

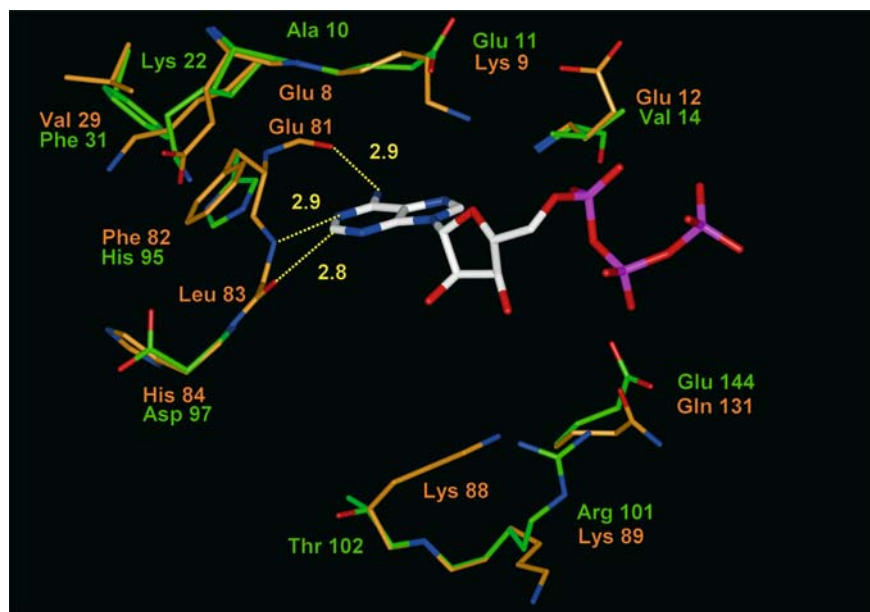


Figure 2. Structural differences between the CDK2 (orange) and CDK4 (green) ATP-binding pockets. For clarity only residues that differ between CDK2 and CDK4 or that interact with the adenine moiety of ATP are shown. The three-pronged hydrogen-bonding pattern (including one aromatic CH...O bond) between ATP and CDK2 [21] is indicated with yellow dotted lines. Hydrogen bond distances are in Å. The backbones of the CDK2 structure and CDK4 model are virtually identical because the CDK2 structure was used as the template for the CDK4 model and there are no insertions or deletions in or near the ATP-binding pocket.

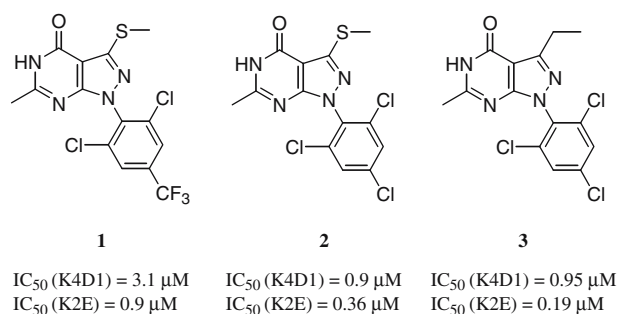


Figure 3. K4D1 and K2E IC₅₀¹ values of the three compounds that led to the potent pyrazolopyrimidinone series: **1** (the initial hit), **2** (a similar, but more active hit), and **3** (the parent compound of the lead series).

lower dielectric of the surrounding medium) than equivalent exposed ones and since the energy minimizations up to this point were carried out with a rigid enzyme, the true binding affinities of both docking modes might be comparable. However, all subsequently synthesized analogs (see Table 2), several of which were specifically designed to discriminate between binding modes, indicate that large substituents in position 6 are well tolerated. This is inconsistent with binding mode A, in which the (6)-methyl group is buried in

the back of the ATP binding site where much larger groups cannot be accommodated. Therefore, we tentatively regarded binding mode B as correct.

In addition to the two expected hydrogen bond interactions, binding mode B has the 3-ethyl substituent located near Phe 93 (CDK2:Phe 80) and pointing in the direction of a buried pocket. In the K2A/ATP structure [21] this pocket (formed by residues 33, 51, 55, 64, 80 and 146) contains two water molecules (167 and 287). Since all residues

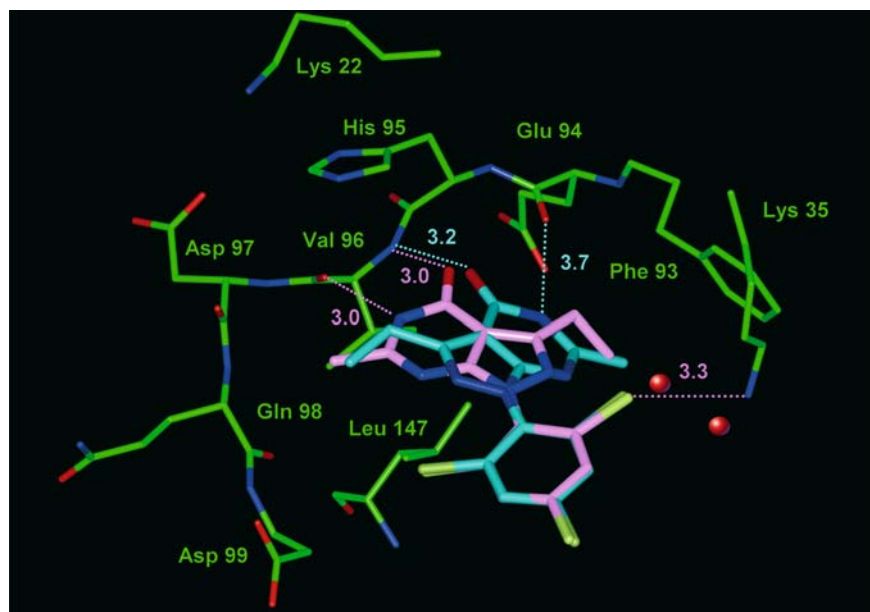


Figure 4. The two binding modes of **3** to CDK4 (mode A: cyan; mode B: pink). Based on the SAR provided by the subsequently synthesized **3** analogs and the crystal structure of **6** complexed with CDK2, binding mode B is most likely correct. Hydrogen bonds are indicated by dotted lines. Distances are in Å. The two red spheres between Lys 35 and the trichlorophenyl group are water molecules taken from the 1FIN crystal structure [21]. They were not energy minimized in the CDK4 model; they are included here only to indicate the location of the CDK4 “water pocket” (formed by Lys 35, Glu 56, Leu 60, Val 72, Phe 93, and Phe 159).

around this largely hydrophobic ‘water pocket’ are conserved between CDK2 and CDK4, we assume that this pocket is very similar in the two enzymes, and a potential target to increase binding affinity. The (1)-trichlorophenyl group of **3** constitutes a third interaction site and is in close contact with Lys 35 (Figure 4). The 1, 3 and 6 positions on **3** are all suitable candidates for substitution and these avenues have indeed been pursued. The synthesis and comprehensive SAR are described elsewhere [28].

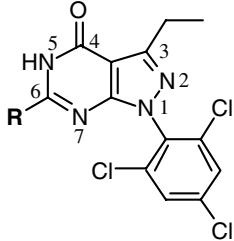
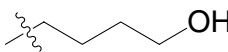
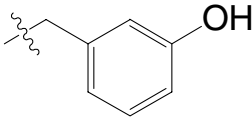
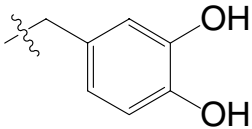
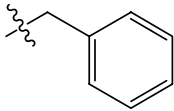
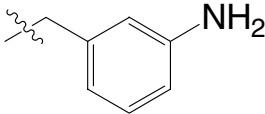
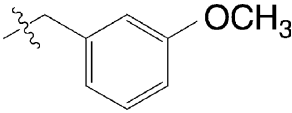
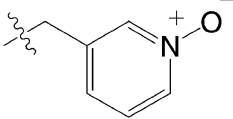
Increasing potency and controlling selectivity through substitutions in position 6 on the pyrimidinone

The binding model of **3** in CDK4 was employed to improve upon the affinity and the specificity for K4D1. Residues were identified that can be exploited specifically by means of polar or charged interactions with substituents on the 6-position of the pyrimidinone. Obvious nearby interaction sites are the sidechains of His 95 and Asp 99, and the backbone carbonyl of Asp 97 in the front of the CDK4 binding pocket. His 95 can also serve as a

selectivity site as this residue is Phe 82 in CDK2, although in the past we had attempted to exploit this difference between CDK2 and CDK4 with other small molecule inhibitors without success. Selectivity may also be obtained through interactions with Lys 22. Although CDK2 has a lysine in the equivalent position (Lys 20), its environment is very different, which may dramatically affect the distribution of its sidechain conformations: Glu 8 and Val 29 in CDK2 keep Lys 20 away from the ATP binding pocket, while Ala 10 and Phe 31 in CDK4 allow Lys 22 to bend towards the binding site (see Figures 2, 5 and 6). Other differences between the CDK2 and CDK4 ATP binding pockets exist (see Table 1), but these concern residues with sidechains that are solvent exposed and/or too far away to interact directly with moderately sized substitutions at the 6-position of the inhibitor.

Aliphatic and aromatic alcohols were designed and synthesized in order to probe the area spanned by His 95 through Asp 99. None of the aliphatic alcohols showed improved affinity. Compound **4**, which was more potent against both enzymes than

Table 2. *In vitro* inhibition of K4D1 and K2E by the parent lead **3** and its analogs. The rightmost column contains the ratio of the K4D1:K2E IC₅₀ values.

				
Number	R	IC ₅₀ nM (K4D1)	IC ₅₀ nM (K2E)	IC ₅₀ ratio
3	H	950	193	4.9
4		2400	250	9.6
5		68	13	5.2
6		44	20	2.2
7		3000	24	125
8		2900	40	72
9		4950	32	124
10		976	79	12

compounds with either longer or shorter chains, is shown as an example in Table 2. In contrast, the K4D1 IC₅₀ values of the phenol **5** and catechol **6** are lower than those of the parent compound **3** by

factors of 14 and 22, respectively. Against K2E, the improvements are 15-fold (**5**) and 10-fold (**6**). Despite the fact that substituents in position 6 are supposed to be directed towards a region that differs between CDK2 and CDK4, the K4D1:K2E IC₅₀ ratios don't differ significantly between **3**, **4**, **5** and **6**. These observations prompt the following questions:

- Can a binding model be produced that explains the strong binding of **5** and **6** and that can qualitatively predict the potency of other meta-substitutions on the (6)-benzyl moiety?
- What functionality is responsible for the improvement of **5** compared to **3** against both enzymes?
- Can the binding model explain the lack of affinity of the aliphatic alcohols?
- Can the lack of selectivity of the most active compounds be explained and what is the prognosis in synthesizing selective ATP competitive inhibitors?

In order to answer these questions, the binding mode of **6** had to be established. Its starting configuration was derived from binding mode B (pink in Figure 4) of compound **3**, with the catechol positioned near Lys 22 and His 95. Energy minimization resulted in the model shown in Figure 5. The core structure of the inhibitor (see top of Table 2) binds similarly to **3**. The meta-hydroxyl group of the catechol serves a dual purpose in accepting a hydrogen bond from Lys 22 and donating one to His 95 (CDK2:Phe 82). Starting with this initial binding mode of **6**, *BForce* [26] simulations were carried out to search conformational space more extensively. This resulted in fifteen binding modes, none of which were lower in energy than the initial binding mode of Figure 5. With two potential hydrogen bond donors and acceptors on the catechol moiety of **6**, however, it cannot be assumed a priori that its binding mode is similar to the one derived for **3**. The only alternative binding mode that features burial of a large part of the inhibitor and hydrogen bonding with the catechol involves rotation of the inhibitor by approximately 180° about an axis that roughly coincides with the pyrazole-pyrimidinone ring juncture. In this orientation (not shown), the catechol is located near Phe 93 and its diol interacts with Lys 35, while the pyrimidinone amide group no longer interacts directly

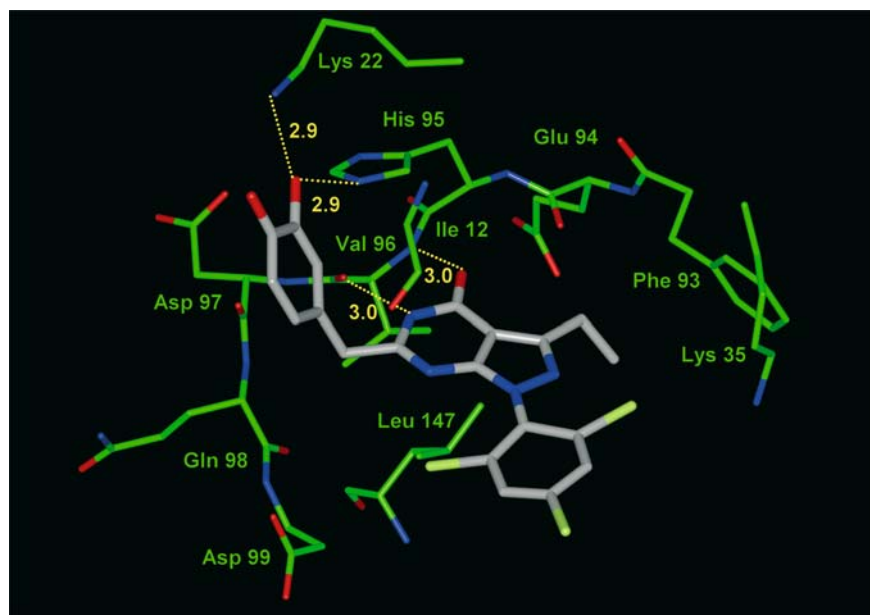


Figure 5. Binding model of **6** to CDK4. Hydrogen bonds are indicated by yellow dotted lines. Distances are in Å.

with the enzyme. *BForce* simulations with this starting configuration led to five new binding modes that were all substantially higher in energy than that of Figure 5.

Concurrent with the computational determination of the binding mode, compounds **7**, **8**, and **9** were synthesized to experimentally determine the critical components of the affinity of **3** and **6** for both enzymes and to validate the binding model. The IC_{50} values of **7** are very surprising, and they indicate that the determinants of K4D1 and K2E affinity are very different: benzyl substitution affects the K4D1 IC_{50} adversely compared to the parent compound (3000 nM for **7**; 950 nM for **3**), but hydroxyl substitution more than compensates for this loss of affinity (68 nM for **5**). The K2E situation is completely reversed: here the benzyl group of **7** is clearly responsible for affinity, while metasubstitution (even with a hydroxyl group, which had such a dramatic effect on K4D1 affinity) has little or no effect. The IC_{50} values of compounds **7**, **8**, **9** and **10** show conclusively that a hydrogen bond donor/acceptor pair in the meta-position of the benzyl group is crucial for affinity to K4D1. N-oxide-containing compound **10** was designed specifically to determine whether a hydrogen bond-donating group in the meta-position is necessary for affinity. Compound **9**

cannot be used to that end as, in addition to lack of hydrogen bond-donating functionality, its methoxy group is bulkier than the substituents in **5** and **8**. Compound **10** is 5× more active than **9**, but 14× less than **5**, so one can conclude that steric effects are less important than the loss of hydrogen bond-donating capabilities. The IC_{50} ratios of **10** and **7** (12 and 125, respectively) indicate that K4D1 has a hydrogen bond donor that K2E lacks. The conserved Lys 22 (CDK2:Lys 20), whose sidechain was postulated to be in different positions in the two enzymes due to different local environments, may well serve this purpose. These observations lend credence both to the CDK4 model and to the binding model of the **3–10** series to CDK4.

Using the **3/6** binding mode, the hydroxyl group of **4** can reach both Asp 99 and His 95. Apparently, neither interaction is favorable. Asp 99 (CDK2:Asp 86) is solvent exposed and little is to be gained in replacing interactions with solvent by interactions with the inhibitor especially when that is associated with substantial conformational entropy loss of the butanol chain. In the previous paragraph it was established that **5** and **6** bind to His 95 and Lys 22. Compound **4** can engage in similar interactions without incurring much intramolecular strain. In doing so, however, it would

pay the same entropic penalty that is associated with binding to Asp 99. Being more rigid, **5** and **6** lose much less conformational entropy.

This rationale for the large effect of meta-substitution on the benzyl group on K4D1 binding but hardly any on K2E binding, does not explain why an unsubstituted benzyl group increases the IC₅₀ ratio from 5 (**3**) to 125 (**7**). Possibly, the local effect of partially desolvating Lys 22 and His 95 without the introduction of compensating hydrogen-bonded interactions contributes to this shift in ratio. This effect would explain the decreased K4D1 affinity, but cannot explain the increased K2E affinity. Although no explicit calculations of lipophilic and electrostatic potentials were carried out to support this, a more probable explanation is that higher free energy penalties are paid for desolvation of the CDK4 ATP binding pocket (with a net -3 charge) upon inhibitor binding than for desolvation of the net neutral CDK2 pocket (see Table 1). This idea is being pursued by augmenting potent inhibitors in several series with additional polar and charged groups to simultaneously improve solubility and K4D1 selectivity. Additionally, the crystal structure of olomoucine and its derivatives in CDK2 show that the benzyl group of olomoucine binds in a similar area of CDK2 [29, 30] as the benzyl substitution on **7**. Olomoucine is known to be a potent inhibitor of CDK2 but not as effective for CDK4, therefore, the addition of the benzyl group may be adding to the potency of the CDK2 inhibition. The increased affinity of **7** vs. **3** with CDK2 may be due to a pi-pi interaction between the benzyl group of **3** and Phe 82 and the CH- π O interaction between the benzyl group and Glu 8.

Final confirmation of the CDK4/4 binding model by the CDK2/4 crystal structure

After the modeling work described in this paper was completed, an in-house crystal structure of the complex of **6** and inactive CDK2 (i.e., not bound to cyclin A or cyclin E) was determined. The experimental details are described elsewhere [28]. The binding of the pyrazolopyrimidinone core to the CDK4 model and to the CDK2 structure (shown in Figure 6) are virtually identical: two hydrogen bonds are formed between Leu 83 and the pyrimidinone ring and the 3-ethyl side chain is located near

CDK2:Phe 80 (CDK4:Phe 93). Three differences between the model and the structure were noted:

- In the model the (1)-trichlorophenyl group makes contact with CDK4:Lys 35, but the crystal structure features a direct interaction with the backbone carbonyl of CDK2:Ile 10. The difference involves a rotation of 90° about the inter-ring bond. When that rotation was applied to compound **6** in the CDK4 model, energy minimization resulted in a binding mode of slightly lower energy. Although the CFF96 forcefield is parameterized for bi-aryl compounds, it appears that the original conformation was a local minimum and a reasonable conformation, but not the lowest energy conformation.
- The pyrazole N2 is not engaged in any direct interaction with CDK4 in the model, but accepts a hydrogen bond from CDK2:Lys 33 in the structure. The position of this lysine in the complex of CDK2 and **6** is significantly different from its position in K2A/ATP [21]. As this is an area where significant conformational changes take place upon cyclin binding, it is possible but not certain that the hydrogen-bonded interaction between CDK2:Lys 33 and the pyrazole N2 would persist in active (cyclin-bound) CDK2.
- The catechol moiety of **6** in the CDK2 structure is rotated about the benzylic bond by 180° from its position in the CDK4 model. In other words, the benzene ring of the catechol is in the same place in the model and the structure, but the meta-hydroxyl groups are at opposite sides of the ring. This difference was fully expected: as indicated before, the meta-hydroxyl group forms hydrogen bonds with His 95 and Lys 22 in the CDK4/compound **6** model, while no such interactions with CDK2 are possible. In the CDK2/compound **6** structure, the meta-hydroxyl group is close to the backbone carbonyl of Ile 10 (see Figure 6), but based on the IC₅₀ data in Table 2, this interaction does not contribute significantly to the K2E affinity of **6**. As a final experiment, the K2E crystallographic binding mode of **6** was minimized in the CDK4 model. In the resulting binding mode, the ligand retains the CDK2 binding motif; the meta-hydroxyl group is in the region of the carbonyl of Ile 10 (Ile 12 in CDK4), but is not within hydrogen bonding distance and the para-hydro-

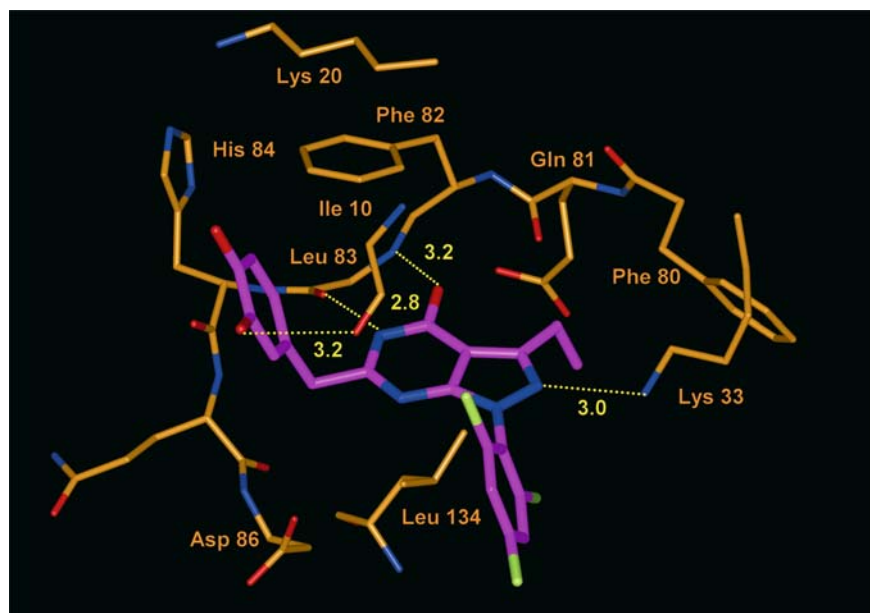


Figure 6. Part of the crystal structure of the complex of **6** with CDK2 showing the binding mode of **6** in the ATP binding pocket. Hydrogen bonds and other relevant nonbonded distances are indicated by yellow dotted lines. Distances are in Å.

xyl remains in the same position as it does seen in Figure 5. Despite conserving the crystallographic conformation, the energy was 7 kcal/mol higher than the predicted binding mode, providing further support for the predicted binding mode. The similarity of the pyrazolo-pyrimidinone core positions between the CDK4/**6** model and the CDK2/**6** structure and the fact that, hydroxyl positions aside, the catechol ring is in the same position in the model and structure provide additional support for the CDK4/**6** binding model.

Employing an iterative cycle of chemistry and modeling, a potent series of K4D1 and K2E inhibitors was designed, synthesized and tested. A CDK4 homology model, based on the 1FIN K2A/ATP structure [21], was effectively used for docking inhibitors, obtaining a single, plausible binding mode, explaining affinity differences within the lead series, and qualitatively predicting the effects of substitutions. The CDK4/**6** binding model was validated by the biological data and by the CDK2/**6** crystal structure. The lack of affinity of the aliphatic alcohols is probably due to entropic differences. Charge differences between the ATP binding pockets in the two enzymes may explain why increased hydrophobicity of inhibitors (e.g., **7**

vs. **3**) leads to tighter CDK2 but not CDK4 binding. Increased hydrophilicity of inhibitors will improve solubility and may in general also lead to better K4D1 over K2E selectivity. Two residues (Lys 22 and His 95 in CDK4, corresponding to Lys 20 and Phe 82 in CDK2) were identified as specific targets for obtaining selectivity for CDK4 over CDK2. The model reveals a cooperative hydrogen bond network between Lys 22, the phenolic hydroxyl group on **5** and **6**, and His 95. Since the position of Lys 20 in CDK2 is probably different from that in CDK4 and Phe 82 in CDK2 cannot accept a hydrogen bond, comparable interactions cannot exist in CDK2. The CDK2/inhibitor crystal structure confirms this. K4D1selective inhibitors were not obtained yet, but based on this work, it seems possible to exploit small differences in the generally very similar ATP binding pockets of CDKs in the ongoing search for selective, ATP competitive, small molecule inhibitors.

Conclusions

Cyclin-dependent kinases (CDKs) and cyclins, their activating agents, play an important role in cell cycle regulation. Naturally occurring macromolecular inhibitors of the CDK family are

involved in tumor suppression. This has prompted efforts to identify small molecule CDK inhibitors for use in cancer chemotherapy. The targets pursued are CDK4/cyclin D1 (K4D1) and CDK2/cyclin E (K2E), which control entry into the DNA replication phase of the cell cycle. Highthroughput screening resulted in the discovery of a low micromolar, ATP-competitive CDK inhibitor with a pyrazolopyrimidinone core. Based on this initial hit, an iterative cycle of chemistry and modeling led to a 70-fold improvement in potency for K4D1 and K2E inhibitors. A CDK4 homology model was built using a CDK2/cyclin A crystal structure as the template. Manual and automated docking of inhibitors into the ATP binding pocket of this CDK4 model suggested a single binding mode. This CDK4/inhibitor binding model is fully consistent with the available SAR and with a subsequently determined CDK2/inhibitor crystal structure. Charge differences between the ATP binding pockets in the two enzymes may explain why predominantly hydrophobic inhibitors have tighter binding to CDK2 than to CDK4. Small changes in substitution patterns resulted in considerable shifts in selectivity for K2E over K4D1 (with an IC₅₀ ratio range from 125 to 2). Although the goal of obtaining K4D1-selective inhibitors has not yet been reached, the modeling did reveal that inhibitor selectivity is largely due to interactions with only two residues (Lys 22 and His 95 in CDK4 and the corresponding Lys 20 and Phe 82 in CDK2). This demonstrates that small differences between enzymes, even in generally similar areas, can be efficiently exploited in the design of selective inhibitors.

Acknowledgments

We thank Melissa J. Ashbacher, Fariba Shoarinejad, Debra A. Doleniak, Marc R. Arnone, Laura M. Handel, Marvin L. Kendall, Robert H. Grafstrom, Xu Xu and Marja M. Dubay for carrying out the *in vitro* studies, and James E. Krywko for building the initial CDK4 models.

References

1. Vermeulen, K., Van Bockstaele, D.R. and Berneman, Z.N., *Cell Prolif.*, 36 (2003) 131.
2. Pines, J., *Nat. Cell Biol.*, 1 (1999) E73.
3. Hung, D.T., Jamison, T.F. and Schreiber, S.L., *Chem. Biol.*, 3 (1996) 623.
4. Rao, R.N., *Curr. Opin. Oncol.*, 8 (1996) 516.
5. McDonald, E.R.I. and El-Deiry, W.S., *Int. J. Oncol.*, 16 (2000) 871.
6. Pardee, A.B., *Science*, 246(4930) (1989) 603.
7. Draetta, G., *Trends Biochem. Sci.*, 15 (1990) 378.
8. Sherr, C.J., *Cell*, 73(6) (1993) 1059.
9. Hall, M. and Peters, G., *Adv. Cancer Res.*, 68 (1996) 67.
10. Jiang, W., Zhang, Y., Kahn, S.M., Hollstein, M.C., Santella, R.M., Lu, S., Harris, C.C., Montesano, R. and Weinstein, I.B., *Proc. Natl. Acad. Sci. USA*, 90 (1993) 9026.
11. Wang, J., Chenivresse, X., Henglein, B. and Brechot, C., *Nature*, 343(6258) (1990) 555.
12. Xiong, Y., Hannon, G.J., Zhang, H., Casso, D., Kobayashi, R. and Beach, D., *Nature*, 366 (1993) 701.
13. Kamb, A., Gruis, N.A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S.V., Stockert, E., Day, R.D. III, Johnson, B.E. and Skolnick, M.H., *Science*, 264(5157) (1994) 436.
14. Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J. and Pavletich, N.P., *Nature*, 382 (1996) 325.
15. Wick, S.T., Dubay, M.M., Imanil, I. and Brizuela, L., *Oncogene*, 11 (1995) 2013.
16. Carlson, B.A., Dubay, M.M., Sausville, E.A., Brizuela, L. and Worland, P.J., *Cancer Res.*, 56 (1996) 2973.
17. 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. and Morishima, H., *J. Med. Chem.*, 44 (2001) 4615.
18. 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. and Morishima, H., *J. Med. Chem.*, 44 (2001) 4628.
19. Ikuta, M., Kamata, K., Fukasawa, K., Honma, T., Machida, T., Hirai, H., Suzuki-Takahashi, I., Hayama, T. and Nishimura, S., *J. Biol. Chem.*, 276 (2001) 27548.
20. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F. Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., *J. Mol. Biol.*, 112 (1977) 535.
21. Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. and Pavletich, N.P., *Nature*, 376(6538) (1995) 313.
22. Sali, A. and Blundell, T.L., *J. Mol. Biol.*, 234 (1993) 779.
23. Sali, A., Potterton, L., Yuan, F., Vlijmen, H. van and Karplus, M., *Proteins Struct. Funct. Genet.*, 23 (1995) 318.
24. Vriend, G. and Sander, C., *J. Appl. Crystallogr.*, 26 (1993) 47.
25. Vriend, G., *J. Mol. Graph.*, 8 (1990) 52.
26. Luty, B.A., Wasserman, Z.R., Stouten, P.F.W., Hodge, C.N., Zacharias, M. and McCammon, J.A., *J. Comput. Chem.*, 16 (1995) 454.
27. Stouten, P.F.W., Froemmel, C., Nakamura, H. and Sander, C., *Mol. Simul.*, 10 (1993) 97.
28. Markwalder, J.A., Arnone, M.R., Benfield, P.A., Boisclair, M., Burton, C.R., Chang, C.-H., Cox, S.S., Czeraniak, P.M., Dean, C.L., Doleniak, D., Grafstrom, R., Harrison, B.A., Kaltenbach, R.F., Nugiel, D.A., Rossi,

- K.A., Sherk, S.R., Sisk, L.M., Stouten, P.F.W., Trainor, G.L., Worland, P. and S.S., P., *J. Med. Chem.*, 47 (2004) 5894.
29. De Azevedo, W.F., Leclerc, S., Meijer, L., Havlicek, L., Strnad, M. and Kim, S.H., *Eur. J. Biochem.*, 243 (1997) 518.
30. Schulze-Gahmen, U., Brandsen, J., Jones, H.D., Morgan, D.O., Meijer, L., Vesely, J. and Kim, S.-H., *Proteins Struct. Funct. Genet.*, 22 (1995) 378.