

S-A-1

**On the Crystal Structure of a human Cell Division Cycle Controlling Protein Kinase(CDK2) and Structure-Based Drug Design**

Sung-Hou Kim

Department of Chemistry and Lawrence Berkeley Laboratory,  
University of California, Berkeley CA 94720, U.S.A.

The most common conventional method of discovering a drug involves a massive screening of a large number of compounds in chemical libraries or in the extracts from natural sources such as plants or microbial broths followed by chemical modification of one or more active compounds to improve their properties as a drug. When the three-dimensional structure of the target molecule for which the drug is searched is known the drug discovery process can be significantly simplified. This is especially true when the three-dimensional structure of a complex between the target and a lead compound is known. In this lecture our experience on the structure-based drug design for human CDK2( cyclin-dependent protein kinase 2) will be discussed with special emphasis on the strength and weakness of this approach of drug discovery. The regulation of the activity of CDK2 plays an important role in the cell proliferation of normal and cancer cells.

**Introduction**

Like other eukaryotic cells, human cell division cycle involves a complex sequence of biochemical events, including the replication of the genome in S phase and the partitioning of cell contents in mitosis. The timing and order of these biochemical events are tightly regulated so that each event in the cycle only occurs after previous events have been completed. Extensive genetic analyses in yeast and biochemical studies in vertebrate cells have recently demonstrated that the timing and coordination of cell cycle events are governed by a class of protein kinases known as CDKs. These kinases undergo periodic activation at various cell cycle stages and are believed to directly trigger the major transitions of the cell cycle (for reviews see Murray and Kirschner 1989; Forsburg and Nurse, 1991; Morgan, 1992; Norbury and Nurse, 1992, Nasmyth and Hunt, 1993). Coordination of cell cycle events is thought to occur at the level of CDK activation: thus, a failure to complete one event (e.g.. DNA replication)

inhibits CDK activities required to initiate the next event (e.g. mitosis). Similarly, cell cycle control by external growth regulatory influences is also thought to involve the regulation of CDK activity. Thus the regulation of CDK activity is of central importance in our understanding of cell division processes of normal and cancer cells and how those processes are regulated by external growth controls.

The activity of a single CDK kinase is required at two transition points in the yeast cell cycle: the commitment to DNA replication in G1 and the induction of mitosis at the G2/M boundary (Forsburg and Nurse 1991). Multiple homologues of these yeast proteins have been identified in human cells. Three of these human CDK-related proteins (known as CDC2, CDK2, and CDK3) are highly related to yeast *cdc2/CDC28* (65% identical) and are capable (to varying extents) of complementing *cdc2* mutants in *S. pombe* or *cdc28* mutants in *S. cerevisiae* (Lee and Nurse 1987; Elledge and Spottswood 1991; Ninomiya-Tsuji et al. 1991; Meyerson et al. 1992). CDC2 was the first higher eukaryotic homologue to be discovered (Lee and Nurse 1987) and has been the most intensely studied. It is most active during mitosis, and has been implicated mainly in the control of the G2/M transition; its role at the G1/S boundary is unclear (Draetta and Beach 1988; Riabowol et al. 1989; Th'ng et al. 1990). Human CDK2 was discovered more recently (Elledge and Spottswood 1991; Ninomiya-Tsuji et al. 1991; Tsai et al. 1991; Rosenblatt et al. 1992). It is most active during S phase and G2 (Rosenblatt et al. 1992).

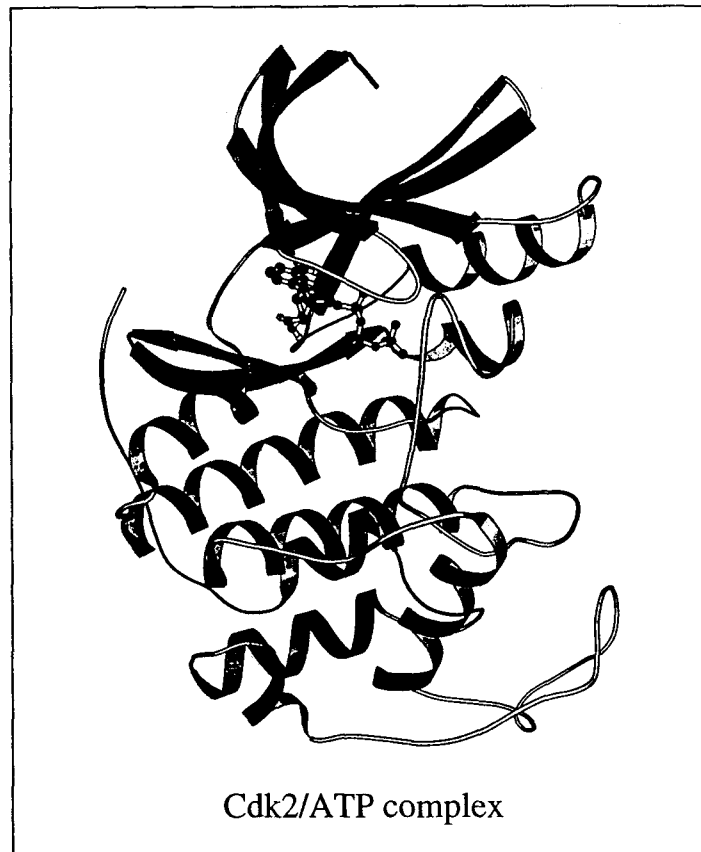
Oscillations in CDK activity during cell cycle are induced by a complex array of post-translational mechanisms. CDKs are small (~34 kD) proteins containing little more than the conserved catalytic core found in all eukaryotic protein kinases (Hanks and Quinn 1991). They are inactive as monomers, and activation requires physical association with regulatory subunits known as cyclins followed by phosphorylation at threonine 160 of CDK2 or 161 of CDC2. Cyclins are a diverse family of proteins (~36 to 50 kD) whose levels oscillate during the cell cycle (thus inducing cell cycle-dependent changes in CDK activity) (Hunt 1989). Deactivation of CDK's are achieved by dephosphorylation of the residue 160/161, phosphorylation of residues 14 or 15, desociation and degradation of bound cyclines, and binding of CDK inhibitor proteins.

In human cells, different cell cycle events appear to be regulated by distinct classes of cyclins. Five classes of human cyclins have been identified (Cyclins A-E) (Pines and Hunter 1989; Pines and Hunter 1990; Lew et al. 1991; Matsushime et al. 1991; Motokura et al. 1991; Xiong et al. 1991). Human cyclin B associates with CDC2 during mitosis to form an active complex involved in the

induction of mitotic events (Pines and Hunter 1989). Cyclin A associates mainly with CDK2 during S phase and G2 and has been implicated in the control of S phase (Girard et al. 1991; Tsai et al. 1991; Elledge et al. 1992; Pagano et al. 1992; Rosenblatt et al. 1992). Cyclin E also associates with CDK2 to form kinase complexes that are active in late G1 and early S phase (Koff et al. 1992). Cyclin C and D have also been implicated in G1 control, but not all of their associated CDK subunits are identified.

Biochemical analysis of CDC2 and CDK2 in vertebrate cells has revealed that the activation of the CDK/cyclin complex is extensively regulated by phosphorylation. Cyclin binding alone does not activate the CDK subunit. Upon binding cyclin, the CDK subunit is phosphorylated on a threonine residue (T161 in CDC2 and T160 in CDK2) by a separate protein kinase known as the CDK activating kinase (Desai et al. 1992; Solomon et al. 1992). Phosphorylation at this site results in the activation of the CDK/cyclin complex (Booher and Beach 1986; Gould et al. 1991; Desai et al. 1992; Gu et al. 1992; Solomon et al. 1992). Thus the mechanism of CDK activation is similar to that of several other protein kinases whose activity depends on phosphorylation at the analogous site in their sequence. For example, activation of cAMP-dependent kinase occurs after autophosphorylation at this position (T197) (Taylor et al. 1990; Knighton et al. 1991). Similarly, MAP kinases are activated when a separate protein kinase (the 'MAP kinase kinase') phosphorylates a threonine and a tyrosine at this site (Payne et al. 1991). CDKs differ from these other kinases in that the activating phosphorylation occurs only after binding of the cyclin regulatory subunit. Thus the substrate binding site of the CDK activating kinase may interact in part with cyclin sequences, conferring cyclin-dependence on the reaction. On the other hand, the binding of cyclin may simply alter CDK conformation to allow T160/1 phosphorylation.

Human CDC2 and CDK2 are also regulated by phosphorylation on tyrosine 15 (Y15) and to a lesser extent on the adjacent threonine residue (threonine 14, T14). These residues are located in the ATP binding site of the kinase, and their phosphorylation inhibits the kinase activity of the T160/1-phosphorylated CDK/cyclin complex (Gould and Nurse 1989; Krek and Nigg 1991; Norbury et al. 1991; Gu et al. 1992; Parker et al. 1992; Solomon et al. 1992). Phosphorylation of these sites occurs after cyclin binding and is probably catalyzed by a separate protein kinase known as wee1 (Parker et al. 1992). Dephosphorylation is catalyzed by a phosphatase known as CDC25 (Dunphy and Kumagai 1991; Forsburg and Nurse 1991; Galaktioniov and Beach 1991; Gautier et al. 1991; Strausfeld et al. 1991; Gu et al. 1992). Phosphorylation of these sites suppresses



The backbone structure of human CDK2 as determined by x-ray crystallographic method(H. DeBonds *et al.*, 1993).  $\beta$ -strands are shown as extended ribbons and  $\alpha$ -helices as coiled ribbons. ATP is represented by a stick model.

the activity of a subpopulation of CDK/cyclin complexes in the cell. Abrupt dephosphorylation of these sites in CDC2 at G2/M contributes to the mitotic surge of CDC2 activity (Gould and Nurse 1989; Krek and Nigg 1991; Norbury *et al.* 1991; Solomon *et al.* 1992). More recently a 21KD protein, p21, has been discovered to be also an inhibitor of the activity of CDK/cyclin ( Xiong *et al.*, 1993; Serrano *et al.*, 1993; Gu *et al.*, 1993; Hunter, 1993).

### Structure of CDK2

The backbone structure of CDK2 can be divided into two lobes (Fig. 1). The smaller N-terminal lobe of CDK2 consists of a sheet of five antiparallel  $\beta$  strands

( $\beta$ 1- $\beta$ 5) and a single large helix ( $\alpha$ 1). The larger C-terminal lobe contains a pseudo four-helical bundle ( $\alpha$ 2, 3, 4, 6), a small  $\beta$ -ribbon ( $\beta$ 6- $\beta$ 8), and two additional helices ( $\alpha$ 5, 7). The ATP binding site is found in the cleft between the two lobes. The core (the  $\beta$ -sheet and the helical bundle) of the CDK2 structure is very similar to that of the cAPK catalytic subunit and MAP kinase, even though these proteins are only distantly related at the amino acid sequence level. This clearly argues that the catalytic cores of all eukaryotic protein kinases exhibit considerable three dimensional structural similarity, as predicted from primary structural homologies at key positions.

***The ATP binding site.*** There already exist several inhibitors of protein kinases that compete for the ATP binding site. Thus ATP binding site is a good target for the structure-based drug design. Our analysis of the crystal structure of CDK2:Mg<sup>++</sup>ATP binary complex indicates that ATP interacts with several residues lining the cleft between the two lobes. The adenine base is positioned in a hydrophobic pocket between the  $\beta$ -sheet of the small lobe and the L7 loop between  $\beta$ 5 and  $\alpha$ 2. The ATP phosphates are held in position by ionic and hydrogen-bonding interactions with several residues, including Lys33, Asp145, and the backbone amides of the glycine-rich loop between  $\beta$ 1 and  $\beta$ 2. An oxygen from each of the three phosphates of ATP contributes to the octahedral coordination of the Mg<sup>++</sup> ion. The three other ligands involved in the coordination of Mg<sup>++</sup> are Asp145, Asn132, and a water molecule. ATP binding appears to induce a slight closure of the cleft by a 2.1 degree hinge movement around an axis parallel to the longitudinal axis of the ATP molecule. Two highly conserved residues (Lys33 and Asn132) exhibit significantly different orientations in the presence and absence of ATP.

***The protein substrate binding site.*** A second potential target site for the structure-based drug design is the protein substrate binding site. Based on the  $\gamma$ -phosphate position of ATP in the crystal structure of CDK2:Mg<sup>++</sup>ATP complex, protein substrates presumably bind in the large cleft between the two lobes, allowing the substrate serine hydroxyl to be positioned near the  $\gamma$ -phosphate of ATP. Interestingly, the core of the protein substrate binding cleft in CDK2 is almost completely blocked by the large loop (the 'T-loop', residues 152-170, containing the Thr160 phosphorylation site) that follows the  $\alpha$ L12 helix. Several residues in this loop block access to the  $\gamma$ -phosphate of ATP, and effective binding of the protein substrate to CDK2 would seem to be impossible. Thus, the 'T-loop' appears to be acting as an auto-inhibitor of protein substrate binding. The quality of the electron density in this region of the CDK2 structure suggests that a portion of this loop (residues 156 to 162) is highly dynamic and possibly flexible

in solution.

### **Structure-Based Drug Design: Crystal Structures of the Complexes between CDK2 and Small Inhibitors**

Most non-protein inhibitors of protein kinases have rather broad specificity: an inhibitor of one protein kinase usually have a significant inhibitory activity for other protein kinases. Thus the initial steps in the structure-based drug discovery process involve biochemical screening of a small set of compounds that are very likely to be good candidates as inhibitors of CDK based on the known biochemical properties or the three-dimensional structure of CDK or CDK-cofactor complex. Since ATP is a cofactor of the protein a series of derivatives of adenine or adenosine have been screened for their inhibitory properties (Rialet and Meijer, 1991). Among those screened two classes of adenine derivatives showed inhibitory activities, one of which being specific to CDKs. We have determined the crystal structures of the complexes of both inhibitors. Based on these structures we learned the strength of the structure-based drug design: The structure revealed the structural basis for the specificity of one of these inhibitors, thus providing a set of specific suggestions about how to improve specificity of this class of inhibitors and some important lessons on common misconception about the structure-based drug design. The weakness of this approach was revealed by the most surprising discovery from these structures that the orientations of the adenine moiety of the two inhibitors and that of ATP are all different. Thus, despite our knowledge of CDK2:ATP structure we could have never guessed the orientation of the adenine moieties of the two drugs based on the structures alone.

### **Acknowledgement**

The work described here has been funded by the U. S. Department of Energy, and done in collaboration with the laboratories of Dr. David Morgan of University of California-San Francisco and Dr. Lawrant Meijer of CNRS-Roscoff, France.

### **References**

- Booher, R. and D. Beach (1986). Mol. Cell. Biol. 6: 3523-3530.
- Desai, D., Y. Gu and D. O. Morgan (1992). Mol. Biol. Cell 3: 571-582.

- Draetta, G. and D. Beach (1988). Cell **54**: 17-26.
- Dunphy, W. G. and A. Kumagai (1991). Cell **67**: 189-196.
- Elledge, S. J., R. Richman, F. L. Hall, R. T. Williams, N. Lodgson and J. W. Harper (1992). Proc. Natl. Acad. Sci. USA **89**: 2907-2911.
- Elledge, S. J. and M. R. Spottswood (1991). EMBO J. **10**: 2653-2659.
- Forsburg, S. L. and P. Nurse (1991). Ann. Rev. Cell Biol. **7**: 227-256.
- Galaktionov, K. and D. Beach (1991). Cell **67**: 1181-1194.
- Gautier, J., M. J. Solomon, R. N. Booher, J. F. Bazan and M. W. Kirschner (1991) Cell **67**: 197-211.
- Girard, F., U. Strausfeld, A. Fernandez and N. J. C. Lamb (1991). Cell **67**: 1169-1179.
- Gould, K. L., S. Moreno, D. J. Owen, S. Sazer and P. Nurse (1991). EMBO J. **10**: 3297-3309.
- Gould, K. L. and P. Nurse (1989). Nature **342**: 39-45.
- Gu, Y., J. Rosenblatt and D. O. Morgan (1992). EMBO J. in press.
- Gu, Y., Turck, C. W., and D. O. Morgan (1993) Nature **366**, 707-710.
- Hanks, S. K. and A. M. Quinn (1991). Meth. Enz. **200**: 38-62.
- Hunt, T. (1989). Curr. Opin. Cell Biol. **1**: 286-274.
- Hunter, T (1993) Cell **75**, 839-841.
- Knighton, D. R., J. Zheng, L. F. Ten Eyck, V. A. Ashford, N.-H. Xuong, S. S. Taylor and J. M. Sowadski (1991). Science **253**: 407-414.
- Koff, A., M. Ohtsubo, A. Giordano, D. Desai, K. Yamashita, J. W. Harper, S. Elledge, T. Nishimoto, D. O. Morgan, R. Franza and J. M. Roberts (1992). Science in press.
- Krek, W. and E. A. Nigg (1991).EMBO J. **10**: 305-316.
- Krek, W. and E. A. Nigg (1991). EMBO J. **10**: 3331-3341.
- Lee, M. G. and P. Nurse (1987). Nature **327**: 31-35.
- Lew, D. J., V. Dulic and S. I. Reed (1991). Cell **66**: 1197-1206.

- Matsushime, H., M. F. Roussel, R. A. Ashmun and C. J. Sherr (1991). Cell **65**: 701-713.
- Morgan, D. O. (1992). Curr. Opin. Genetics Develop. **2**: 33-37.
- Motokura, T., T. Bloom, H. G. Kim, H. Juppner, J. V. Ruderman, H. M. Kronenberg and A. Arnold (1991). Nature **350**: 512-515.
- Murray, A. W. and M. W. Kirschner (1989). Science **246**: 614-621.
- Nasmyth, K. and T. Hunt (1993) Nature, **366**, 634-635.
- Ninomiya-Tsuji, J., S. Nomoto, H. Yasuda, S. I. Reed and K. Matsumoto (1991). Proc. Natl. Acad. Sci. USA **88**: 9006-9010.
- Norbury, C., J. Blow and P. Nurse (1991). EMBO J. **10**: 3321-3329.
- Norbury, C. and P. Nurse (1992). Ann. Rev. Biochem. **61**: 441-470.
- Pagano, M., R. Pepperkok, F. Verde, W. Ansorge and G. Draetta (1992). EMBO J. **11**: 961-971.
- Parker, L. L., S. Atherton-Fessler and H. Piwnicka-Worms (1992). Proc. Natl. Acad. Sci. USA **89**: 2917-2921.
- Payne, D. M., A. J. Rossomando, P. Martino, A. K. Erickson, J.-H. Her, J. Shabanowitz, D. F. Hunt, M. J. Weber and T. W. Sturgill (1991). EMBO J. **10**: 885-892.
- Pines, J. and T. Hunter (1989). Cell **58**: 833-846.
- Pines, J. and T. Hunter (1990). Nature **346**: 760-763.
- Riabowol, K., G. Draetta, L. Brizuela, D. Vandre and D. Beach (1989). Cell **57**: 393-401.
- Rialet, V. and L. Meijer (1991). Anticancer Research **11**, 1581-1590.
- Rosenblatt, J., Y. Gu and D. O. Morgan (1992). Proc. Natl. Acad. Sci. USA **89**: 2824-2828.
- Serrano, M., G. J. Hannon, and D. Beach Nature **366**, 704-707
- Solomon, M. J., T. Lee and M. W. Kirschner (1992). Mol. Biol. Cell **3**: 13-27.
- Strausfeld, U., J. C. Labbé, D. Fesquet, J. C. Cavadore, A. Picard, K. Sadhu, P. Russell and M. Dorée (1991). Nature **351**: 242-245.
- Taylor, S. S., J. A. Buechler and W. Yonemoto (1990). Ann. Rev. Biochem. **59**: 971-1005.



Th'ng, J. P. H., P. S. Wright, J. Hamaguchi, M. G. Lee, C. J. Norbury, P. Nurse and E. M. Bradbury (1990). Cell **63**: 313-324.

Tsai, L.-H., E. Harlow and M. Meyerson (1991). Nature **353**: 174-177.

Xiong, Y., T. Connolly, B. Futcher and D. Beach (1991). Cell **65**: 691-699.

Xiong, Y., G. J. Hannon, H. Zhang, D. Casso, R. Kobayashi, and D. Beach (1993) **366**, 701-704