**Research Article** [Krithiga & Jayachitra, 3(9):Sep., 2012]
**CODEN (USA): IJPLCP** **ISSN: 0976-7126**

INTERNATIONAL JOURNAL OF PHARMACY & LIFE SCIENCES

# Molecular docking studies between cdk and kaempferol from *Clitoria ternatea*

**N. Krithiga and A. Jayachitra***
Department of Plant Biotechnology, School of Biotechnology, Madurai Kamaraj University, Madurai, (TN) - India

## Abstract

In the present study, it has been found out that the CDK protein which is present in the entire organism originated from the common ancestral algae (Chlamydomonas *reinhardtii* and *Volvox carteri f. nagariensis).* In this work, an attempt has been made to find out the interaction of CDK protein with which is the essential polyphenols present in the extract of *Clitoria ternatea* are compared. From the observation obtained in his study, it is concluded that kaempferol has the ability to interact with CDK protein with he minimum energy.

Key-Words: Kaempferol, *Clitoria ternatea*, Molecular docking

## Introduction

Kaempferol is a strong antioxidant and helps to prevent oxidative damage of our cells, lipids and DNA. Kaempferol seems to prevent arteriosclerosis by inhibiting the oxidation of low density lipoprotein and the formation of platelets in the blood. Studies have also confirmed that kaempferol acts as a chemopreventive agent, which means that it inhibits the formation of cancer cells.

**Cell cycle**

Cell division is one of the most conspicuous features of life, and thus several elements of the control of cell division are common to both prokaryotes and eukaryotes (Amon, 1998; Leatherwood, 1998). The degree of evolutionary conservation is especially striking among eukaryotes, where progression through the successive phases of the cell cycle (S, G2, M, and G1) in species as diverse as yeast and humans is driven by a common class of heterodimeric serine/threonine protein kinases. These kinases consist of a catalytic subunit, termed Cyclin-dependent Kinase (CDK), and an activating subunit, cyclin (Nigg, 1995).

**Cyclin dependent Kinase**

CDKs are the catalytic subunits of a large family of heterodimeric serine/threonine protein kinases whose best-characterized members are involved in controlling progression through the cell cycle.

*** Corresponding Author**
E.mail: jchitra21@gmail.com
Ph. No.: +0452-2458273

According to the latest versions of the human and mouse genomes, there are 11 genes encoding CDKs and 9 other genes encoding CDK-like proteins with conserved primary structure. Because the catalytic activity of these CDKs requires binding of a regulatory subunit, the term CDK is often used for the active heterodimeric complex. The activating partners of the cell-cycle CDKs, are molecules that are synthesized and degraded during each cell cycle and thus have been designated 'Cyclins'. Although this property has been used to define this Kinase family, not all activating partners of CDKs are synthesized and destroyed in a cyclical fashion. Indeed, the physiological role of most CDKs and their activating partners remains unknown.

**Activation of cyclin dependent Kinase**

Activation of CDKs also requires phosphorylation of conserved Thr residues within their T-loops, such as residues Thr161 and Thr160 in human CDK2. This activating phosphorylation is catalyzed by CDK-activating kinases (CAKs) (Kaldis, 1999). Upon phosphorylation of Thr160 by CAK, the T-loop moves away and the Cyclin A-CDK2 complex becomes fully active (Jeffrey *et al*., 1995; Russo *et al*., 1996).

**G1 Cyclin-dependent kinases**

In mammalian cells, there are two classes of CDKs that function at the G1/S phase transition (Sherr, 1994; Reed, 1997; Draetta, 1994). CDK4 and its close relative CDK6 are driven by three D-type Cyclins: D1, D2 and D3. The primary target of these activities is pRb and related proteins (Grana, 1998). Cyclin E accumulates very close to the G1/S phase transition and specifically activates CDK2. Although Cyclin-E –

CDK2 has a secondary role in phosphorylating pRb (Zhang *et al*., 2000; Harbour *et al*., 1999), after Cyclin-D – Cdk4/6 the critical target(s) of Cyclin-E – CDK2 in regulation of S phase is not known.

**Regulation of Cyclin - dependent Kinase activity by Cyclins**

Cyclin abundance oscillates during the cell cycle as a result of programmed synthesis and degradation, thereby assuring a limited window of CDK activation. Over expression of D-type Cyclins or Cyclin E during early G1 leads to premature S phase entry (Ohtsubo *et al*., 1993; Quelle *et al*., 1993; Resnitzky *et al*., 1994) suggesting that the G1 Cyclins are at least partially rate limiting for S phase entry and confirming that regulation of CDK activity by Cyclin accumulation has biological consequences.

**Cyclin-dependent Kinase function**

Even in yeast, using sophisticated genetic analysis, only a few key targets have been identified. This is probably due to the output of CDK function being achieved by the phosphorylation of many proteins and not by the activation of linear signal cascades that are easy to identify both genetically and biochemically. One notable exception is the mammalian retinoblastomas susceptibility protein, pRb. pRb is a component of a transcriptional repression module that targets many genes whose products function during S phase or at the G1/S transition. Phosphorylation of pRb by G1 CDKs near the G1/S boundary relieves this repressive activity and allows transcription of S phase genes (Zhang *et al*., 2000; Harbour *et al*., 1999).

**Phylogenetic Tree**

A phylogenetic tree or evolutionary tree is a branching diagram or "tree" showing the inferred evolutionary relationships among various biological a species or other entities based upon similarities and differences in their physical and/or genetic characteristics. The taxa are joined together in the tree are implied to have descended from a common ancestor.

**GOR method**

The GOR method (**G**arnier-**O**sguthorpe-**R**obson) is an information theory -based method for the prediction of secondary structure in proteins. The GOR method analyzes sequences to predict alpha helix, betasheet, turn, or random coil secondary structure at each position based on 17-amino-acid sequence windows. The four matrices reflect the probabilities of the central, ninth amino acid being in a helical, sheet, turn, or coil conformation. In subsequent revisions to the method, the turn matrix was eliminated due to the high variability of sequences in turn regions (particularly over such a large window). The method was considered as best requiring at least four contiguous residues to score as alpha helices to classify the region as helical, and at least two contiguous residues for a beta sheet.

**Docking**

In the molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to in turn predict the affinity and activity of the small molecule. Two approaches use a matching technique that describes the protein and the ligand as complementary surfaces. The second approach stimulates the actual docking process in which the ligand – protein pair wise interaction energies are calculated. Docking program depends on the two components: the search algorithm and scoring function. Hence docking plays an important role in the rational design of drugs.

## Material and Methods

**Sequence retrieval**

Many bioinformatics processes require a sequence as input. Retrieval of sequence is an important step in any bioinformatics analysis. Each and every sequence has its own unique accession number in any database. Sequence can be retrieved either by querying with keywords or by the accession number.

**NCBI Database**

The sequences used in this study were retrieved from NCBI database (View: http://www.ncbi.nlm.nih.gov/). The National Center for Biotechnology Information (NCBI) is part of the United Stated National Library of Medicine (NLM), a branch of the National Institutes of Health. The NCBI is located in Betherda, Maryland and was founded in 1988 through legislation sponsored by Senator Claude Pepper. The NCBI houses genome sequencing data in Gen bank and an index of biomedical research articles in Pubmed Central and Pubmed, as well as other information relevant to biotechnology. All these databases are available online through the Entrez search engine.

**Phylogenetic Tree**

The sequence retrieved from NCBI database were multiple aligned and then the phylogenetic tree was constructed. A phylogenetic tree or evolutionary tree is a branching diagram or "tree" showing the inferred evolutionary relationships among various biological a species or other entities based upon similarities and differences in their physical and/or genetic characteristics. The taxa joined together in the

tree are implied to have descended from a common ancestor.

**GOR method**

The GOR method (**G**arnier-**O**sguthorpe-**R**obson) is an information theory -based method for the prediction of secondary structure in proteins. It was developed in the late 1970s shortly after the simpler Chou-Fasman method. The GOR method analyzes sequences to predict alpha helix, beta sheet, turn, or random coil secondary structure at each position based on 17-amino-acid sequence windows. The original description of the method included four scoring matrices of size 17×20, where the columns correspond to the log-odds score, which reflects the probability of finding a given amino acid at each position in the 17-residue sequence. The four matrices reflect the probabilities of the central, ninth amino acid being in a helical, sheet, turn, or coil conformation. In subsequent revisions to the method, the turn matrix was eliminated due to the high variability of sequences in turn regions (particularly over such a large window). The method was considered as best requiring at least four contiguous residues to score as alpha helices to classify the region as helical, and at least two contiguous residues for a beta sheet.

**Pubchem**

Pubchem is a database of chemical molecules and their activities against biological assays. The system is maintained by the National Center for Biotechnology Information (NCBI), a component of the National Library of Medicine, which is part of the United States National Institutes of Health (NIH). Millions of compound structures and descriptive datasets can be freely downloaded via FTP. Pubchem contains substance descriptions and small molecules with fewer than 1000 atoms and 1000 bonds.

**Protein data bank**

The protein data bank (PDB) is a repository for the 3D structural data of large biological molecules, such as proteins and nucleic acids. The data, typically obtained by X-ray crystallography or NMR spectroscopy and submitted by biologists and biochemists from around the world, can be accesses at no charge on the internet. The PDB is overseen by an organization called the world wide Protein Data Bank.

**Molecular docking**

The interaction between the main antioxidant phytochemical found in the clitoria ternatea Kaempherol with the CDK 2 proteins was studied using online molecular docking server package

## Results and Discussion

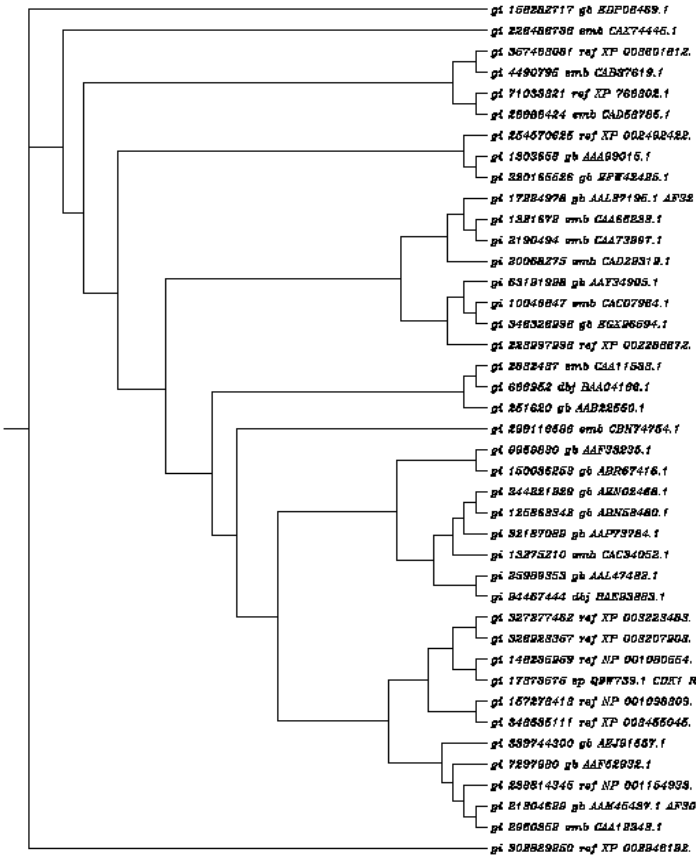


**Fig. 1: Phylogenetic analysis of CDK protein among various organism (Fungi, Plants, animals, Protozoa and Metazoa**

**GOR4 result for: Chlamydomonas reinhardtii**

GOR secondary structure prediction method version IV, J. Garnier, J.-F. Gibrat, B. Robson, Methods in Enzymology,R.F. Doolittle Ed., vol 266, 540-553, (1996)

View GOR4

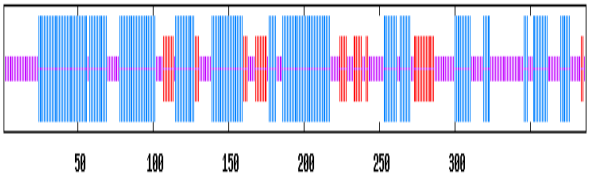
```
      10        20        30        40        50        60        70
       |         |         |         |         |         |         |
MAPGFGNFATADDASTSSGYQDQGPLARLLSKL
RQFKALAADKDLANSELAPLIRALPLDTELQQLL
AKF
ccccccccccccccccccccchhhhhhhhhhhhhhhhhhhhhhhh
hhhhhhhhhhchhhhhhhhhhh
HWYPGHNFTTTVDLANLKGALQKYKYIKIGQLG
SGSYGVVHKAINRETNELLAIKKVVHSIENGLSD
STI
ccccccchhhhhhhhhhhhhhhhhhhhhhhhcccccceeeeeech
hhhhhhhhhhhheeeccccccch
REISTLRELQHDNIVRLKDIIATVNGTHVHLVLEF
LDCDLRHYLDTYAEASNINRIKSIVFQILRGIRHA
```

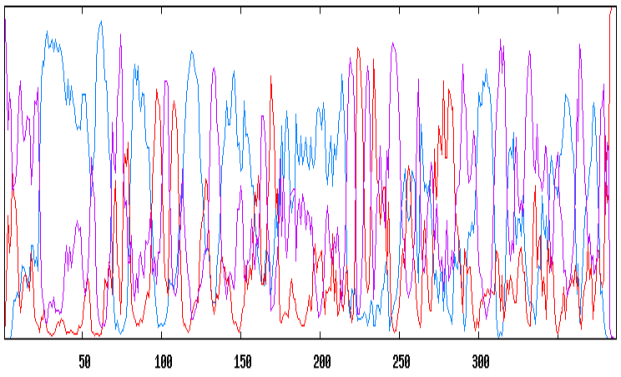
hhhhhhhhhhhhhhhhhhhheeeccccceeeeeeeechhhhhcccc
hhhhhhhhhhhhhhhhhhhhhhhh
HANSIMHRDLKPQNVLVGVHSGNVKITDFGLAR
CFLPNEDRAYTERVVTLYYRAPELLLGAQHYTS
AVDL
hhhhhhhhcccccceeeeeccccceeeeeeccceeccccccccchhhhh
hhhhcchhhhhhhcceeeeeee
WSVGCIMAEMVNFEPLFRSDSEIGLLFRMFEQLG
TPTPDAWHELSGLAYYSENFPRFVPRRFEDMVPR
LA
eeeeeeeccccccccccccchhhhhhhhhhhhccccccchhhhhcccc
ccccccccccccccccchhhc
NDAVGLDLLRRMLCYDPRQRITASEALVHPWFN
DVVV
cchhhhhhhhhhhccccccchhhhhhhhccccccceec

**Sequence length : 387**

GOR4 :

Alpha helix (Hh) : 192 is 49.61%
$3_{10}$ helix (Gg) : 0 is 0.00%
Pi helix (Ii) : 0 is 0.00%
Beta bridge (Bb) : 0 is 0.00%
Extended strand (Ee) : 50 is 12.92%
Beta turn (Tt) : 0 is 0.00%
Bend region (Ss) : 0 is 0.00%
Random coil (Cc) : 145 is 37.47%
Ambigous states (?) : 0 is 0.00%
Other states : 0 is 0.00%




Prediction results file (text): [GOR4]

**Fig. 2: Secondary structure prediction of Chlamydomonas *reinhardtii* by using GORIV method**

**GOR4 result for: Volvox carteri f. nagariensis**

secondary structure prediction method version IV, J. Garnier, J.-F. Gibrat, B. Robson, Methods in Enzymology,R.F. Doolittle Ed., vol 266, 540-553, (1996)

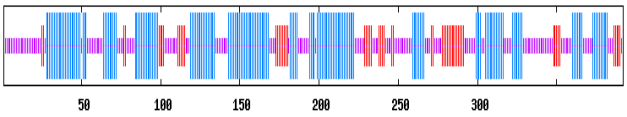
View GOR4

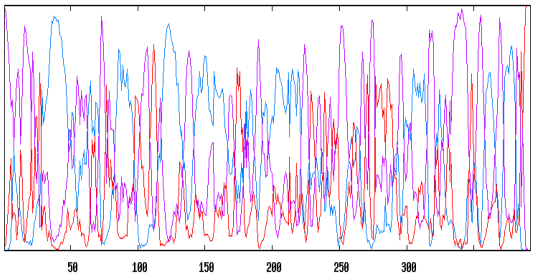
10 20 30 40 50 60 70
| | | | | | |
MAGGHMHAPAANGNILNPGATQYPPITHLLWHL
SHFKALAAERELANVAIAPLVRQLPFDQASNGEL
VSL
ccccccccccccccccccccccceechhhhhhhhhhhhhhhhhhh
hhhchhhccccccccchhhhhh
LEKFGWHPIHDFTTTLDLAGLKGALQKYKYIKM
GPLGSGSYGVVYKAQNRETQELLAIKRVRFSIAE
HGL
hhhcccceeccccchhhhhhhhhhhhhhheeeeccccccceeeeecc
chhhhhhhhhhhhhhhhccccc
SDSTIREISTLRELRHDNIVKLKDIIATANGQHVH
LVLEFLDCDLRQYLDTHPEASEIGRIKSVVLQILR
ccchhhhhhhhhhhhhhhhhhhhhhhhhhhhcccceeeeeeeechhhh
hccccccchhhhchhhhhhhhhh
GIRHAHMNSIMHRDLKPQNVLIGVAKGQVKLTD
FGLARCFLPNSDRAYTERVVTLYYRAPELLLGSP
CYT
hhhhhhhhhhhhhccccccceeeeecccceeeeccccceeccccccccc
chhhhhhhhccccceeccccce
SAVDLWSVGCIMAEMINFEPLFKADTEIGLLFRIF
EKLGTPNLEVWKDLRGLTHFSDDFPNFPPKPMR
QL
eeeeeeeeeeeccccccchhhhcchhhhhhhhhhhhcccccchhhhh
hhccccccccccccccccccee
VPRLAGDPAGLDLLSRLLTYDPSRRITARQALEH
PWFQGVVV
eeeccccccchhhhhhhcccccchhhhhhhhhhcccceeeeec

**Sequence length: 392**

GOR4 :

Alpha helix (Hh) : 172 is 43.88%
$3_{10}$ helix (Gg) : 0 is 0.00%
Pi helix (Ii) : 0 is 0.00%
Beta bridge (Bb) : 0 is 0.00%
Extended strand (Ee) : 58 is 14.80%
Beta turn (Tt) : 0 is 0.00%
Bend region (Ss) : 0 is 0.00%
Random coil (Cc) : 162 is 41.33%
Ambigous states (?) : 0 is 0.00%
Other states : 0 is 0.00%






Prediction result files (text): [GOR4]

**Fig. 3: Secondary structure prediction of *Volvox carteri f. nagariensis* by using GORIV**

Figure 1 shows the phylogenetic tree for CDK sequence for plant and animal kingdom. Among those 41 organisms were selected and its sequence were retrieved from NCBI database. To construct a phylogenetic tree the sequence was first multiple aligned using FASTA tool. Then the phylogenetic tree was constructed using clustalw software. The result shows Chlamydomonas *reinhardtii* and *Volvox carteri f. nagariensis are* the ancestral organisms for this protei

Figure 2 and 3 shows the secondary structure of Chlamydomonas *reinhardtii* and *Volvox carteri f. nagariensis* were predicted by using the GOR IV method. In this the sequence length of *Chlamydomonas reinhardtii* is 387. Then the secondary stucture predicted result shows that it contained 49.61%, alpha helix region, 12.92% of Extended strand, 37.47% of random coil and 310 helix, pi helix, beta bridge, beta turn bend region, Ambigous states and other states were found to be absent and for *Volvox carteri f. nagariensis* possess the sequence length is 392. It contained 43.88% of alpha helix region, 14.8%of extended strand and 41.33% of random coil.

But the 310 helix, pi helix Beta bridge, beta turn, Bend region, Ambigous states and other states shows 0% result which indicates their absence in the secondary structure prediction

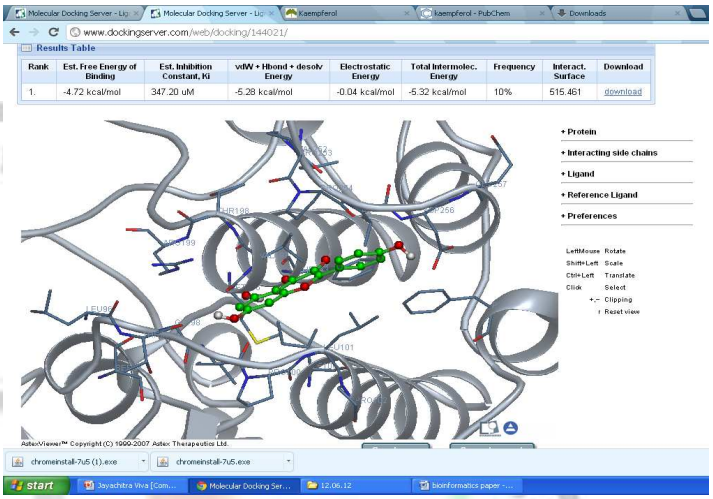



**Fig. 4: Interaction of CDK2 protein with kaempherol compound in *Clitoria ternatea***

Docking of kaempherol with CDK2 protein is shown in the figure 4. Based on the information available it has been shown clearly that kaempherol compound in *Clitoria ternatea* has been used as ligand for the CDK2 protein. Kaempherol compound on interaction with CDK2 protein produced an bond energy value -5.28 Kcal/mol, Estimated free energy of binding -4.72kcal/mol, the interaction surface area of 515.461 and the estimated inhibition constant (ki) 347.20 µM. Active groups present in the CDK2 protein were the site of binding to the ligand EGCG. The energy values were calculated using by using molecular modeling software how the pharmacophoric part of the drug was initially identified.

The phylogenetic tree was constructed for both plant and animal kingdom of CDK family. The evolutionary rates of the plant kingdom and animal kingdom for CDK family members were expected to be different. Therefore phylogenetic analysis was performed by the NJ method (Felsenstein, 1981; Saitou and Nei, 1987). These methods are less likely to mislead than parsimony or compatibility method when the evolutionary rates differ among taxa (Felsenstein, 1978; Saitou and Imanishi, 1989). When taxa have dissimilar rates of substitutions, incorporating the differential rate of character changes at given sites by fitting the amino acid data set to gamma rate distributions facilitates a more accurate phylogeny (Yang, 1996).

The specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners (Canduri *et al*., 2001; De Azevedo, 2001; Kim, 1996). CDK2 crystal

structures identified to date shows that ligands binding at the ATP binding- site should form a hydrogen bond. Based on the information available in the literature it has been shown clearly that kaempherol in *Clitoria ternatea* has been used to interact with CDK2 protein.

## References

1. Amon A, (1998). Controlling cell cycle and cell fate: Common strategies in prokaryotes and eukaryotes. *Proc. Natl. Acad. Sci*., USA. 95**:** 85–86.
2. Canduri F, Teodoro L.G.V.L., Lorenzi C.C.B, Hial V, Gomes R.A.S, Ruggiero J, Neto W.F and De Azevedo Jr, (2001). Crystal structure of human uropepsin at 2.45*_AA resolution*. *Acta Crystallogr.*, D 57: 1560–1570.
3. De Azevedo W.F. Jr., Canduri F, Fadel V, Teodoro L.G.V.L., Hial V and Gomes, R.A.S, (2001). Molecular model for the binary complex of uropepsin and pepstatin, *Biochem. Biophys. Res. Commun*., 287 (1): 277–281.
4. Draetta GF, (1994). Mammalian G1 cyclins**.** *Curr. Opin. Cell Biol.*, 6: 842-846.
5. Draetta G, (1987). Identification of p34 and p13, human homologs of the cell cycle regulators of fission yeast encoded by cdc2C and suc1C. *Cell*, 50: 319–325.
6. Felsenstein, (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol*., 17: 368–376.
7. Felsenstein J, (1978). Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool*., 24: 401–410.
8. Harbour JW, Luo RX, Dei Santi A, Postigo AA, Dean DC, (1999). Cdk ·· phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1**.** *Cell*, 98: 859-869.
9. Jeffrey P.D, Russo A.A, Polyak K, Gibbs E, Hurwitz J, Massague J, and Pavletich N.P, (1995). Mechanism of CDK activation revealed by the structure of a cyclin A-CDK2 complex. *Nature*, 376: 313–320.
10. Kaldis P, (1999). The cdk-activating kinase (CAK): From yeast to mammals. *Cell. Mol. Life Sci*., 55: 284–296.
11. Leatherwood J**,** (1998). Emerging mechanisms of eukaryotic DNA replication initiation. *Curr. Opin. Cell Biol*., 10**:** 742–748.
12. Ohtsubo M and Roberts JM, (1993). Cyclin-dependent regulation of G1 in mammalian fibroblasts**.** *Science*, 259: 1908-1912.
13. Resnitzky D, Gossen M, Bujard H and Reed SI, (1994). Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. *Mol Cell Biol* , 14: 1669-1679.
14. Russo A.A, Jeffrey P.D., and Pavletich N.P, (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat. Struct. Biol*., 3: 696–700.
15. Saitou N and Imaishi T, (1989). Relative efficiencies of the Fitch-Margoliash, maximum parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. *Mol. Biol. Evol.*, 6: 514–525.
16. Saitou N, and Nei M, (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol.Biol*. Evol., 4: 406–425.
17. Sherr CJ, (1994). G1 phase progression: cycling on cue**.** *Cell*, 79**:** 551-555.
18. Yang Z, (1996). Among site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol*, 11: 367–372.
19. Zhang HS, Gavin M, Dahiya A, Postigo AA, Ma D, Luo RX, Harbour JW, Dean DC, (2000). Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF**.** *Cell*, 101: 79-89.