The EMBO Journal vol.10 no.9 pp.2653-2659, 1991

A new human p34 protein kinase, CDK2, identified by complementation of a *cdc28* mutation in *Saccharomyces cerevisiae*, is a homolog of *Xenopus* Eg1

Stephen J.Elledge and Matthew R.Spottswood

Department of Biochemistry and Institute for Molecular Genetics, Baylor College of Medicine, Houston, TX 77030, USA

Communicated by V.Pirrotta

The onset of S-phase and M-phase in both *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* requires the function of the *cdc2/CDC28* gene product, p34, a serine-threonine protein kinase. A human homolog, $p34^{cdc2}$, was identified by functional complementation of the *S.pombe cdc2* mutation (Lee and Nurse, 1987). Using a human cDNA expression library to search for suppressors of *cdc28* mutations in *S.cerevisiae*, we have identified a second functional p34 homolog, *CDK2* cell division kinase). This gene is expressed as a 2.1 kb transcript encoding a polypeptide of 298 amino acids. This protein retains nearly all of the amino acids highly conserved among previously identified p34 homologs from other species, but is considerably divergent from all previous $p34^{cdc2}$ homologs, ~65% identity. This gene encodes the human homolog of the *Xenopus Eg1* gene, sharing 89% amino acid identity, and defines a second sub-family of CDC2 homologs. A second chromosomal mutation which arose spontaneously was required to allow complementation of the *cdc28-4* mutation by *CDK2*. This mutation blocked the ability of this strain to mate. These results suggest that the machinery controlling the human cell cycle is more complex than that for fission and budding yeast.

Key words: CDC complementation/Eg1/CDK2/$p34^{cdc2}$

Introduction

In the last several years it has become increasingly clear that the basic mechanisms which regulate the cell cycle have been conserved throughout the evolution of eukaryotes, from yeast to humans. Of particular importance is the finding that a number of the proteins controlling these mechanisms are interchangeable from one organism to another. One of these proteins, the CDC28/cdc2 protein kinase, plays a critical role in cell division control in eukaryotes. In the yeast cell cycle, the induction of both DNA synthesis (S-phase) and mitosis (M-phase) requires this serine-threonine-specific protein kinase which is the product of the *CDC28* gene in *Saccharomyces cerevisiae* (Hartwell *et al.*, 1974; Reed, 1980; Lorincz and Reed, 1984) and the $cdc2^+$ gene in *Schizosaccharomyces pombe* (Nurse and Bissett, 1981). It has been suggested that $p34^{cdc2}$ performs these mutually exclusive functions by phosphorylating different sets of substrates through a cell cycle phase-dependent association with different proteins that alter its substrate specificity and/or catalytic efficiency (Pines and Hunter, 1990b). These proteins are known as cyclins and their abundance varies through the cell cycle (Evans *et al.*, 1983; Swenson *et al.*, 1986; Standart *et al.*, 1987; Nash *et al.*, 1988; Westendorf *et al.*, 1989; Pondaven *et al.*, 1990; Wittenberg *et al.*, 1990).

In higher eukaryotes, an essential role for the kinase activity has been demonstrated at the G_2 to M-phase transition by several criteria. Microinjection of antibodies against a human homolog, $p34^{cdc2Hs}$, arrests cells in G_2 (Riabowol *et al.*, 1989), a temperature sensitive mutation in the mammalian *CDC2Hs* gene arrests cells at the G_2-M phase boundary at the non-permissive temperature and this arrest can be suppressed by expression of the wild-type human protein (Th'ng *et al.*, 1990). Further supporting evidence for a role in G_2 is that the *CDC2Hs* gene is cell cycle regulated at the level of transcript abundance, being maximally expressed in G_2 (McGowan *et al.*, 1990). In addition, the H1 kinase activity of the CDC2Hs containing complex is maximally active in G_2 (Draetta and Beach, 1988). Although a role in G_2 has been demonstrated for $p34^{cdc2}$ in mammalian cells, its role in G_1 has yet to be clearly demonstrated. It is possible that this protein is specific for G_2 and that other p34 homologs exist which have roles in other stages of the cell cycle. This work discusses the identification of a new human *cdc2* homolog suggesting that the regulation of the human cell cycle is more complex than the cycle of yeast.

Results

Isolation of human cDNAs capable of suppressing the conditional lethal phenotype of the cdc28-4 mutation in S.cerevisiae

A number of cell cycle regulatory genes have been cloned in yeast by overproduction suppression of mutant phenotypes. For example, *CLN1*, *CLN2* and *CKS1* were originally identified as dosage-dependent suppressors of *cdc28* mutations in *S.cerevisiae* (Hadwiger *et al.*, 1989a,b). The $suc1^+$ gene from *S.pombe* was identified in a similar manner as a suppressor of certain *cdc2* alleles (Hayles *et al.*, 1986). We sought to exploit the evolutionary conservation of function among cell cycle regulatory proteins to identify mammalian proteins capable of suppressing temperature-sensitive *cdc28* mutations. A human cDNA library was constructed in the yeast cDNA expression vector λ YES-R (Elledge *et al.*, 1991) from mRNA prepared from EBV transformed human peripheral lymphocytes. λ YES-R is a cDNA expression vector containing the *URA3*, *ARS1* and *CEN4* sequences for selection and replication in yeast and *bla* and pBR322 *ori* sequences for replication and selection in *Escherichia coli*. cDNAs inserted into λ YES-R come under the control of the *GAL1* promoter. 1×10^7 independent recombinants were obtained, amplified, and converted into plasmids by *cre*-mediated automatic excision as described. This library was transformed into the yeast strain Y10 (*cdc28-4*), pooled, and grown at 30°C in liquid

minimal media lacking uracil with glucose as a carbon source. Approximately 10^6 independent transformants were obtained in this experiment. Cells were plated on YP galactose plates, incubated at 30°C for 5 h to induce the *GAL1* promoter, and then transferred to the non-permissive temperature, 35°C, and incubated for 3 days. Twenty-four colonies which grew at 35°C were tested for plasmid dependence of the suppression phenotype by measuring their ability to grow on complete minimal media supplemented with 5-fluoro-orotic acid (5-FOA) at 30°C and 35°C. 5-FOA selects against cells expressing the *URA3* gene (Boeke *et al.*, 1988). Since minichromosomes containing centromeres are not as stable as larger chromosomes, a fraction of cells in each colony, ~5%, have lost the plasmid and can grow on media containing 5-FOA. Twenty clones were able to grow on 5-FOA at 30°C but not at 35°C and were chosen for further study.

These clones fell into two general classes based on growth rates. Class 1 had 18 members and grew at rates similar to wild-type CDC28 strains. Class 2 had two members and grew slowly at 35°C, but equally well as class 1 clones at 30°C on galactose. Representatives of these classes are shown in Figure 1 grown under selective and non-selective conditions. Both classes showed plasmid dependence for growth at 35°C and galactose dependence, indicating that the suppressing function was under the control of the *GAL1* promoter which is repressed by the presence of glucose. Four class 1 clones were not galactose dependent for growth at 35°C. Plasmid DNAs were recovered from the yeast clones by transformation into *E.coli*. Restriction mapping coupled with DNA hybridization analysis revealed that the clones fell into the two classes which precisely corresponded to the growth rate classes. Class 1 clones had insert sizes ranging from 1.0 kb to 1.6 kb. Both class 2 members each had 2.0 kb inserts. These classes did not show cross hybridization at high stringency. A class 1 clone with a 1.6 kb insert, pSE999, and a class 2 clone, pSE1000, were chosen for further study.

DNA sequence analysis reveals that class 2 cDNAs encode a new human p34 homolog

*Xho*I adaptors were added to the cDNAs as part of the library generation method and neither pSE999 nor pSE1000 contained an internal *Xho*I site. The entire cDNA from each plasmid was introduced into pBS KS^+ and pBS KS^- as *Xho*I fragments and sequenced by the generation of nested deletions as described previously (Elledge and Davis, 1987). Since it was known that the human p34 homolog identified

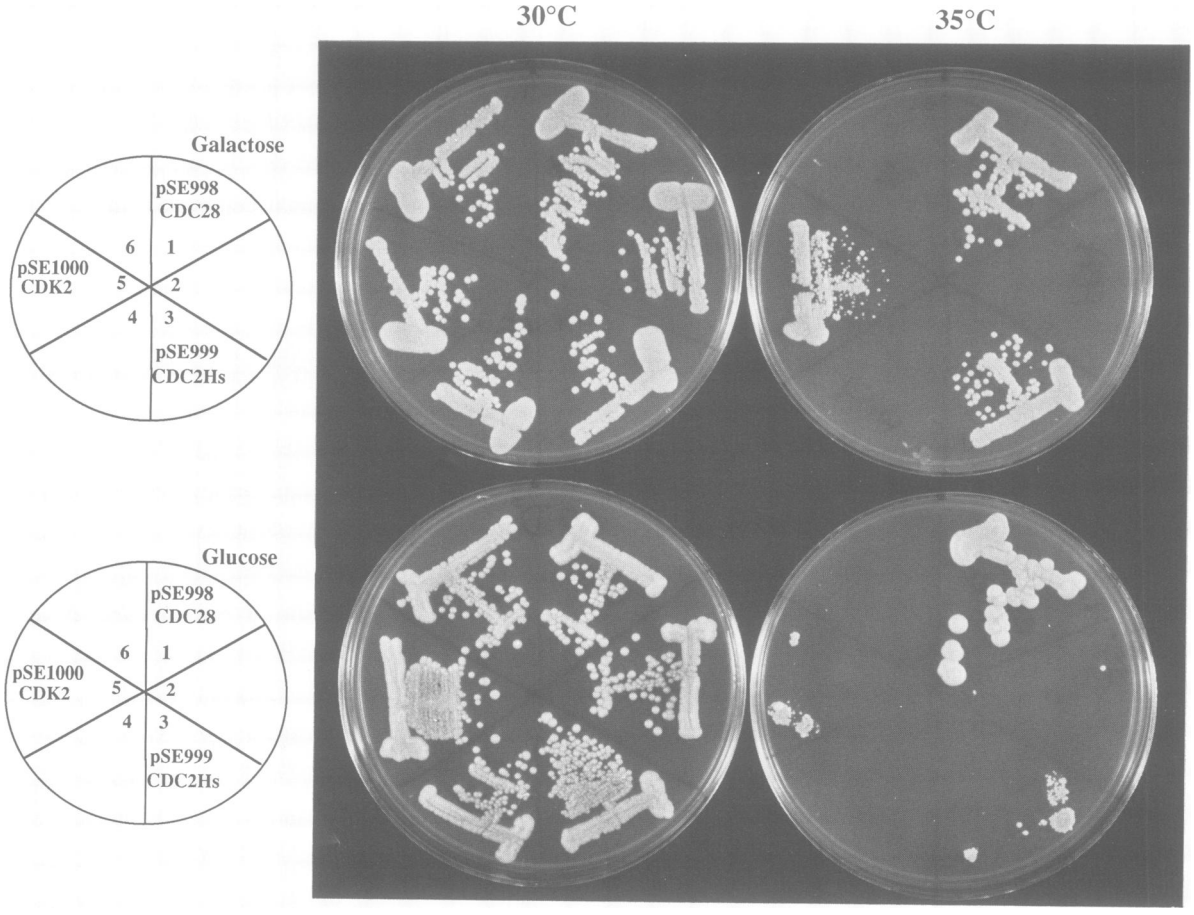

Fig. 1. Complementation of the *cdc28*-4 mutation by pSE998 (*CDC28*), pSE999 (*CDC2Hs*) and pSE1000 (*CDK2*). The key to the left of the plates is a guide to the location of the strains on the plates to the right. Position 1, Y10 containing pSE998, a clone of the *S.cerevisiae CDC28* gene; position 2, Y10 alone; position 3, Y60 containing pSE999 (*CDC2Hs*); position 4, Y60 alone; position 5, Y61 containing pSE1000 (*CDK2*); position 6, Y61 alone. The strains lacking a plasmid were derived from the strain containing the plasmid by growth on media containing 5-FOA. The carbon source for each row is indicated to the left of the row. Other essential nutrients were provided by yeast extract and peptone. The temperature at which each column of plates was grown is indicated at the top of the column.

by Lee and Nurse (1987) through complementation of the *S.pombe cdc2* mutation would also complement the *S. cerevisiae cdc28* mutation (Wittenberg and Reed, 1989), we anticipated isolation of this class of cDNA insert. DNA sequencing analysis revealed that the class 1 cDNA was identical to the $p34^{cdc2}$ gene as described previously (Elledge *et al.*, 1991) and will not be discussed further here. The nucleotide sequence of class 2 cDNAs (Figure 2) predicts a protein of 298 amino acids. This protein shows significant homology to p34 proteins isolated from a number of organisms (Figure 3). Long stretches of amino acid identity are shared among all members of this gene family. In particular the new human homolog shares identity in the ATP-binding site in the amino-terminal region, residues 11–20, and in the EGVPSTAIRELLKE region, residues 42–56. This second region has significance because anti-peptide antibodies to this region have been used frequently to identify the human p34 protein. The most divergent portion of these homologs is located in the central portion of these proteins, amino acids 64–117, and the carboxy-terminal portion beyond amino acid 243. We have named this protein homolog $p34^{cdk2}$ and the gene *CDK2* (cell division kinase) to distinguish it from the homolog discovered by Lee and Nurse (1987).

Amino acid sequence alignment shows that p34 proteins now fall into at least two sub-families based on sequence identity. One sub-family consists of the original CDC2Hs, and the proteins from *Drosophila*, mouse and chicken. The second consists of CDK2, *Xenopus* Eg1 (Paris *et al.*, 1991), and possibly the *Drosophila* 2c protein (Lehner and O'Farrell, 1990). CDK2 and Eg1 are clearly in the same group being 89% identical (Table I), suggesting a possible functional difference with the CDC2Hs class. Placement of the *Drosophila* 2c protein into this class is less clear because although it is more homologous to CDK2 (63%) and Eg1 (65%) than to CDC2Hs (57%), it shows no more overall similarity to the CDK2 class than members of the CDC2Hs class do to the CDK2 class (~64%). One fact suggesting that 2c may be a member of the CDK2 class is that there are 17 amino acid positions at which CDK2, Eg1 and 2c are completely conserved and different from the CDC2Hs class, and one position where an amino acid is deleted, although two are deleted in 2c and only one in Eg1 and CDK2. The isolation of additional homologs from other organisms will allow the more precise definition of CDC2 kinase subclasses.

Expression of CDK2 in human cell lines

It was important to demonstrate that CDK2 was of human origin and not a result of library contamination. Thus, to identify the message in other cell lines of human origin and to determine the message size and potential complexity of the *CDK2* gene, RNA blot analysis was performed. $Poly(A)^+$ mRNA was isolated from Hep 3B2, a human liver cell line (Aden *et al.*, 1979), fractionated on a 1% agarose–formaldehyde gel, blotted to nitrocellulose and hybridized with radioactive *CDK2* and *CDC2Hs* cDNAs at high stringency (Figure 4). A unique 2.1 kb transcript was

```
ATGGAGAACTTCCAAAAGGTGGAAAAGATCGGAGAGGGCACGTACGGAGTTGTGTACAAAGCCAGAAACAAGTTGACGGG 80
 M  E  N  F  Q  K  V  E  K  I  G  E  G  T  Y  G  V  V  Y  K  A  R  N  K  L  T  G 27

AGAGGTGGTGGCGCTTAAGAAAATCCGCCTGGACACTGAGACTGAGGGTGTGCCCAGTACTGCCATCCGAGAGATCTCTC 160
  E  V  V  A  L  K  K  I  R  L  D  T  E  T  E  G  V  P  S  T  A  I  R  E  I  S  53

TGCTTAAGGAGCTTAACCATCCTAATATTGTCAAGCTGCTGGATGTCATTCACACAGAAAATAAACTCTACCTGGTTTTT 240
L  L  K  E  L  N  H  P  N  I  V  K  L  L  D  V  I  H  T  E  N  K  L  Y  L  V  F  80

GAATTTCTGCACCAAGATCTCAAGAAATTCATGGATGCCTCTGCTCTCACTGGCATTCCTCTTCCCCTCATCAAGAGCTA 320
 E  F  L  H  Q  D  L  K  K  F  M  D  A  S  A  L  T  G  I  P  L  P  L  I  K  S  Y 107

TCTGTTCCAGCTGCTCCAGGGCCTAGCTTTCTGCCATTCTCATCGGGTCCTCCACCGAGACCTTAAACCTCAGAATCTGC 400
  L  F  Q  L  L  Q  G  L  A  F  C  H  S  H  R  V  L  H  R  D  L  K  P  Q  N  L  133

TTATTAACACAGAGGGGGCCATCAAGCTAGCAGACTTTGGACTAGCCAGAGCTTTTGGAGTCCCTGTTCGTACTTACACC 480
L  I  N  T  E  G  A  I  K  L  A  D  F  G  L  A  R  A  F  G  V  P  V  R  T  Y  T  160

CATGAGGTGGTGACCCTGTGGTACCGAGCTCCTGAAATCCTCCTGGGCTCGAAATATTATTCCACAGCTGTGGACATCTG 560
 H  E  V  V  T  L  W  Y  R  A  P  E  I  L  L  G  S  K  Y  Y  S  T  A  V  D  I  W 187

GAGCCTGGGCTGCATCTTTGCTGAGATGGTGACTCGCCGGGCCCTGTTCCCTGGAGATTCTGAGATTGACCAGCTCTTCC 640
  S  L  G  C  I  F  A  E  M  V  T  R  R  A  L  F  P  G  D  S  E  I  D  Q  L  F  213

GGATCTTTCGGACTCTGGGGACCCCAGATGAGGTGGTGTGGCCAGGAGTTACTTCTATGCCTGATTACAAGCCAAGTTTC 720
R  I  F  R  T  L  G  T  P  D  E  V  V  W  P  G  V  T  S  M  P  D  Y  K  P  S  F  240

CCCAAGTGGGCCCGGCAAGATTTTAGTAAAGTTGTACCTCCCCTGGATGAAGATGGACGGAGCTTGTTATCGCAAATGCT 800
 P  K  W  A  R  Q  D  F  S  K  V  V  P  P  L  D  E  D  G  R  S  L  L  S  Q  M  L 267

GCACTACGACCCTAACAAGCGGATTTCGGCCAAGGCAGCCCTGGCTCACCCTTTCTTCCAGGATGTGACCAAGCCAGTAC 880
  H  Y  D  P  N  K  R  I  S  A  K  A  A  L  A  H  P  F  F  Q  D  V  T  K  P  V  293

CCCATCTTCGACTCTGATAGCCTTCTTGAAGCCCCCGACCCTAATCGGCTCACCCTCTCCTCCAGTGTGGGCTTGACCAG 960
P  H  L  R  L  *  * 298

CTTGGCCTTGGGCTATTTGGACTCAGGTGGGCCCTCTGAACTTGCCTTAAACACTCACCTTCTAGTCTTAACCAGCCAAC 1040

TCTGGGAATACAGGGGTGAAAGGGGGGAACCAGTGAAAATGAAAGGAAGTTTCAGTATTAGATGCACTTAAGTTAGCCTC 1120

CACCACCCTTTCCCCCTTCTCTTAGTTATTGCTGAAGAGGGTTGGTATAAAAATAATTTTAAAAAAGCCTTCCTACACGT 1200

TAGATTTGCCGTACCAATCTCTGAATGCCCCATAATTATTATTTCCAGTGTTTGGGATGACCAGGATCCCAAGCCTCCTG 1280

CTGCCACAATGTTTATAAAGGCCAAATGATAGCGGGGGCTAAGTTGGTGCTTTTGAGAATTAAGTAAAACAAAACCACTG 1360

GGAGGAGTCTATTTTAAAGAATTCGGTTAAAAAATAGATCCAATCAGTTTATACCCTAGTTAGTGTTTTCCTCACCTAAT 1440

AGGCTGGGAGACTGAAGACTCAGCCCGGGTGGGGGT
```

Fig. 2. The *CDK2* cDNA sequence and deduced amino acid sequence. The deduced amino acid sequence (298 amino acids) of the putative protein is shown in single letter code below the cDNA sequence. The sequence shows the first 1476 nucleotides from the ATG, the remaining 600 nucleotides of the 3′ untranslated region were not sequenced in their entirety and are not shown.

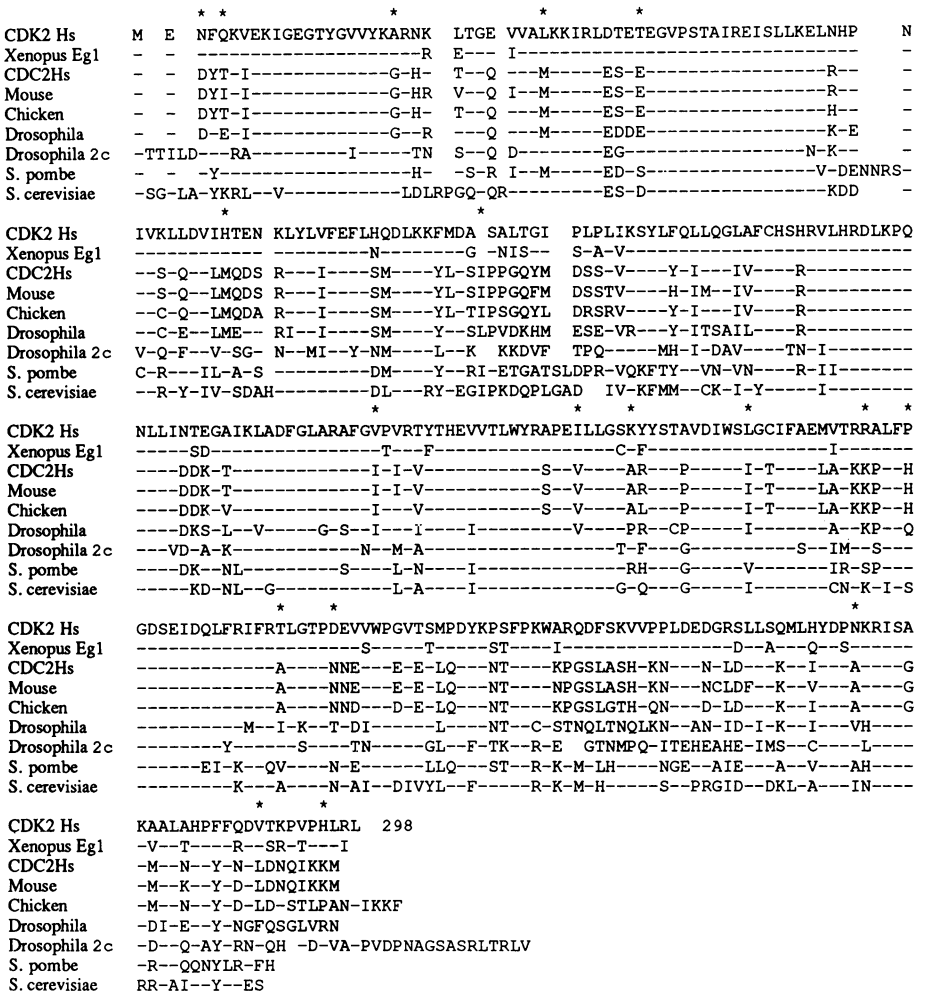

Fig. 3. Comparison of the amino acid sequence of the human CDK2 protein with $p34^{cdc2}$ homologs from humans (Lee and Nurse, 1987), *Xenopus* (Paris *et al.*, 1991), mouse (Spurr *et al.*, 1990), chicken (Krek and Nigg, 1989), *Drosophila* (Jimenez *et al.*, 1990; Lehner and O'Farrell, 1990), *S.pombe* (Hindley and Phear, 1984) and *S.cerevisiae* (Lorincz and Reed, 1984). Sequences were aligned by eye for maximal homology with the minimum number of gaps. Dashes represent identities, spaces in the sequence indicate positions of insertions or deletions in one of the homologs.

observed for CDK2, roughly corresponding to the size of the cDNA isolated by complementation. Two mRNA species were apparent for CDC2Hs in accordance with published reports. Three other human cell lines were tested for CDK2 expression: CEM cells, a B-cell line, MOLT-4, a T-cell line and HeLa cells. All four cell lines gave an mRNA species of 2.1 kb (data not shown). This confirms that *CDK2* is of human origin and is derived from a 2.1 kb polyadenylated transcript. Furthermore it is expressed in all human cell lines tested.

A spontaneous mutation in Y10 allows CDK2 to complement cdc28 mutations

Complementation was shown to be plasmid dependent as described above. It was anticipated that upon reintroduction of the plasmids into Y10, complementation of the *cdc28*-4 mutation would be observed. However, this was not the case. Class 1 clones showed very weak complementation, and class 2 clones completely failed to complement. To investigate this further, strains isolated by complementation were grown on 5-FOA at the permissive temperature to select for isolates that had lost the pSE999 and pSE1000 clones. Both these strains, Y60 and Y61 respectively, became uracil auxotrophs and temperature sensitive for growth, restoring the properties of the parent strain Y10. Introduction of pSE999 or pSE1000 into either of these strains could rescue the temperature sensitivity phenotype when grown with galactose as a carbon source but failed to rescue when glucose was used as a carbon source. Furthermore, pSE999 showed more complete

Table I. Percentage identity between p34 homologs

Gene	% identity to CDK2	% identity to CDC2Hs
CDK2	100	66
CDC2Hs	66	100
Xenopus Eg1	89	64
Mouse	67	96
Chicken	65	93
Drosophila	63	72
Drosophila 2c	65	57
S.pombe	66	66
S.cerevisiae	65	66

Percentage identities were calculated by subtracting the number of amino acid substitutions according to the alignment in Figure 3 from the total number of amino acids in common between the two proteins, ignoring internal amino acid gaps and carboxy-terminal length differences.

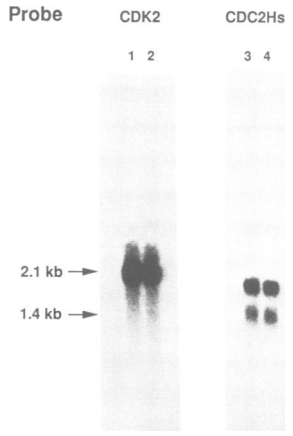

Fig. 4. Expression of CDC2Hs and CDK2. Poly(A)$^+$ RNAs from Hep 3B2 cells were probed on Northern blots with a *CDK2* probe (lanes 1 and 2) or *CDC2Hs* probe (lanes 3 and 4) as indicated. Lanes 1 and 2 each contain 5 μg of poly(A)$^+$ Hep 3B2 RNA from two different preparations of Hep 3B2 RNA. After probing with *CDK2*, the filter was stripped and subsequently probed with a *CDC2Hs* probe.

complementation of the *cdc28*-4 allele in Y60 and Y61 than in Y10.

One possible explanation for the difference between the ability of pSE999 and pSE1000 to complement Y10 could be the expression levels of the clones. The adaptor method employed for construction of the cDNA libraries used for complementation generated a small inverted repeat of 19 bp at the ends of 50% of the cDNAs (Elledge *et al.*, 1991) that may alter the translational context of particular clones. cDNAs from both pSE999 and pSE1000 contained this 19 bp repeat and have differing amounts of 5′ untranslated regions. Either or both of these could contribute to differential translation of these clones. To minimize potential translational differences due to differing translational contexts, polymerase chain reaction (Saiki *et al.*, 1988) was used to create versions of these clones that removed the adaptor and provided identical 5′ untranslated regions. These clones showed properties identical to pSE999 and pSE1000 with respect to complementation of Y10 and Y61 (data not shown), suggesting that the complementation difference was not due to a variation in the translational context of the two messages. However, the new version of *CDK2*, pSE1037, appeared to provide better complementation of Y61 as measured by larger colony size at the non-permissive temperature. Furthermore, the amount of mRNA from cells bearing each construct grown on galactose was quantified and found to be equal, ruling out differential RNA expression or stability (data not shown).

The mutation allowing *CDK2* to complement *cdc28*-4 is likely to have arisen spontaneously in the culture prior to the introduction of the library because multiple independent transformants had acquired the ability to be complemented by *CDK2*. To determine whether the mutations in these strains were linked to *cdc28*-4 or were dominant or recessive, we attempted to mate Y60 and Y61 with different haploids to produce diploids for genetic analysis. However, both strains were sterile, failing to mate with strains of either mating type. They also failed to sporulate on sporulation media suggesting that they are not diploids. The sterility was not dependent upon the *cdc28*-4 mutation, because introduction of a plasmid carrying *CDC28*, pSE998, failed to restore the ability to mate. Further analysis will be required to determine the nature of the suppressing mutation(s) and whether the same mutation is responsible for both the suppression and sterility.

CDC2Hs and CDK2 can complement a null mutation in CDC28

A second possible explanation for the differential ability of pSE999 and pSE1000 to complement the *cdc28*-4 mutation is that CDK2 cannot compete efficiently with the mutant *cdc28*-4 protein for a common subunit of the kinase. In this scenario, the mutant protein encoded by *cdc28*-4 would block the ability of CDK2 to complement in a dominant negative fashion. To test this hypothesis, we assessed the ability of each human kinase to complement a *cdc28* null mutation. A diploid, Y62, was constructed that was heterozygous for an insertion in *CDC28, cdc28::TRP1*. pSE999 and pSE1000 were introduced into Y62 and subsequently sporulated. Tetrads were dissected and allowed to germinate on YP galactose medium at 30°C. *CDC2Hs* was able to complement the *cdc28* null mutation; however, *CDK2* was not (Table II), suggesting that if a dominant negative relationship did exist, it was not the sole block to complementation.

Twice, in the analysis of positions that failed to grow into colonies from the tetrad analysis of Y62 pSE1000, a microcolony of ~100 cells would appear. Micromanipulation of cells from these microcolonies to unique positions revealed that ~50% of these cells would double once approximately every 30 h. We were unable to establish a strain from these cells that would grow faster to analyze the plasmid content of these cells, although at least one of the sister spore clones in each tetrad that generated these microcolonies was a uracil auxotroph suggesting that there was a possibility that these cells contained the pSE1000 plasmid. Microcolonies of this nature were not observed from the tetrad analysis of Y62 alone.

The presence of microcolonies suggested the possibility of weak complementation. Since pSE1037 showed a slightly better complementation of *cdc28*-4 than pSE1000, further tetrad analysis was performed using pSE1037 in the Y62 background. In general, pSE1037 failed to show complementation at the rate anticipated for co-segregation of the plasmid with the *cdc28::TRP1* allele. However, we recovered two tetrads out of 20 in which three colonies grew. In these tetrads, two colonies were large, and one was small. The small colonies were Trp$^+$, galactose dependent, and sensitive to 5-FOA indicating that the cells contained pSE1037 and that it was essential for growth. Southern analysis confirmed that the strains contained the *cdc28::TRP1* allele and lacked the wild-type *CDC28* gene (data not shown). The fact that complementation of the null allele of *CDC28* did not occur at the rate anticipated may reflect a failure to efficiently complement the germination role of *CDC28*, or perhaps complementation depended on a spore receiving two copies of the plasmid instead of one, a rarer event. Although the *CDK2* gene on pSE1037 was derived from PCR and may have a PCR induced mutation, we think that the difference between it and pSE1000's ability to complement the null mutation is more likely to be due to differences in expression levels because *CDK2* driven by a *GAL* promoter on a 2-micron plasmid vector was found to

Table II. Complementation of *cdc28::TRP1* null mutants by *CDC2Hs* and *CDK2*

Strain	Number of viable spores per tetrad	Number of tetrads observed	Genotype			
			trp1ura3	*TRP1URA3*	*trp1URA3*	*TRP1ura3*
Y62 (pSE1000)	4	0	0	0	0	0
(*CDK2*)	3	0	0	0	0	0
	2	19	20	0	18	0
	1	2	1	0	1	0
Y62 (pSE999)	4	6	12	12	0	0
(*CDC2Hs*)	3	13	14	12	13	0
	2	5	1	1	8	0
	1	0	0	0	0	0

Y62 (*CDC28/cdc28::TRP1*) strains bearing either pSE999 (CDC2Hs) or pSE1000 (CDK2) were sporulated. Tetrad analysis was performed on the meiotic products and spores were allowed to germinate on YP media with 21% galactose as a carbon source. All strains recovered with the *TRP1URA3* genotype were found to die on media containing 5-FOA, indicating that the plasmid in these strains is essential for viability.

complement a *cdc28* null (J.Tsuji and K.Matsumoto, personal communication).

In an attempt to isolate further suppressors of *CDK2* that would allow complementation of *cdc28*-4 alleles, we introduced pSE1000 into a *cdc28*-4 background, Y82, with genetic markers more amenable to later genetic analysis than Y10. Surprisingly, *CDK2* was capable of complementing the temperature sensitivity of this strain on galactose media, unlike Y10. This result suggests that the strain background may have a profound influence on the ability to observe complementation.

Discussion

The primary observations of this study are that human cells contain at least two functional *cdc2*-like genes, each of which is capable of complementing some or all of the defects associated with the *cdc28*-4 mutation. Because *CDK2* can perform nearly all of the functions of the CDC28 protein in budding yeast it is reasonable to assume that it performs a similar role to *CDC28* in controlling the human cell cycle. This conclusion is supported by the structural similarity between the genes both in homology and size of the proteins. The presence of this protein in mammalian cells has yet to be demonstrated but will be tested once specific antisera are generated. However, evidence for a second *cdc2*-like protein has appeared in the literature. Pines and Hunter have shown that immunoprecipitation of cyclin A also brings down a 33 kDa phosphoprotein that cross-reacts with antipeptide antibodies directed against the PSTAIRE region of CDC2Hs, but which is not recognized by antisera directed against the carboxy terminus of CDC2Hs (Pines and Hunter, 1990a). Furthermore, this protein has a very similar fragmentation pattern to $p34^{cdc2Hs}$ when cleaved with *N*-chlorosuccinimide which cleaves after tryptophan residues, but a distinct pattern when treated with trypsin which cleaves after arginine and lysine residues. It is possible that *CDK2* encodes this protein because it has a similar size, shares conservation of all four tryptophan residues, has a distinct trypsin fragmentation pattern, and has a divergent carboxy terminus that would not be recognized by antipeptide antisera directed against the carboxy terminus of CDC2Hs.

What are the roles of these *cdc2*$^+$ homologs in the human cell cycle? The most likely points of action during the human cell cycle are analogous to start and the control of mitosis in *S.cerevisiae* and *S.pombe*, in late G_1 at the R-point and in late G_2 at the beginning of mitosis. Studies of the function of *CDC2Hs* in mammalian cells point to a role in mitosis. A *ts* mutation in the *CDC2Hs* gene in the FT210 cell line gives a G_2 arrest at the non-permissive temperature. The CDC2Hs transcript shows cell cycle regulation being maximally expressed in G_2 and gives further, though circumstantial, support for a role in G_2. Interestingly, however, although the message fluctuates, the protein level remains constant. This may be due to the inability to distinguish between the various homologs. Evidence for a role in mitosis has been generated by antibody injection studies. There is also evidence for a role in S-phase demonstrated *in vitro* in oocyte extracts by depletion studies (Blow and Nurse, 1990); however, these experiments cannot rule out a role for indirect inhibition of DNA synthesis by depletion of a cdc2 associated factor, not the cdc2 protein itself. However, based on the genetics in yeast and the relatedness of eukaryotes, it is quite likely that these proteins will have roles in G_1 and G_2 transition points. The question then becomes why do human cells have multiple *cdc2*-like genes and what are their precise functional roles. Are they expressed in different tissues or at different times in development, or at different times in the cell cycle? Do they associate with different proteins to control the phosphorylation of different substrates? Do they control different transition points? It is likely that they do not have completely overlapping functions because *ts* mutants in the *CDC2Hs* gene which give a conditional mitotic arrest phenotype can be overcome by the *CDC2Hs* gene (Th'ng *et al.*, 1990). Assuming that a functional CDK2 protein is expressed in that cell line, it would not be capable of carrying out all of the functions of the *CDC2Hs* gene. This argument is dependent upon the assumption that the *CDC2Hs* mutation is completely recessive. It is important to point out that although there is a clear functional difference between CDC2Hs and CDK2 when expressed in *S.cerevisiae*, this cannot be taken as evidence for functional differences in mammalian cells. Conversely, although CDC2Hs alone can carry out all of the functions of *cdc2*$^+$ and *CDC28* genes in yeast, this does not mean that it is all that is needed in mammalian cells. Clearly it will be important to determine the precise individual roles of these proteins in the mammalian cell cycle. It is likely that Eg1 and CDK2 define a new family of cdc2 kinases and their close sequence similarity suggests a common and distinct cellular function relative to CDC2Hs.

Due to the homologs overall relatedness, many

experiments that rely on polyclonal antisera to CDC2Hs or antipeptide antisera to the PSTAIRE region are likely to cross-react with CDK2 and therefore cannot be used to distinguish among these proteins. We have observed cross-reaction between antisera directed against CDC2Hs with the CDK2 protein (S.J.Elledge and W.Harper, unpublished results). Therefore, many of the properties previously assigned to CDC2Hs must be re-examined when more specific reagents become available.

Materials and methods

Bacterial, yeast and plasmid strains

E.coli JM107 (*end*A1, *gyr*A96,*thi*,*hsd*R17,*sup*E44, *rel*A1,Δ(*lac-pro*AB), {F', *tra*D36,*pro*AB$^+$,*lac*I^{q}ZΔM15) (Yannisch-Perron *et al.*, 1985) was used as a transformation recipient for all plasmid constructions unless indicated otherwise. pMC9 (*lac*I) was a gift from M.Calos. CMY478 (*cdc28-4, ura3-52, try1, ile1, trp1-1*) was a gift from C.Mann. Y10 was an ilv$^+$ revertant of CMY478. pSE998, a genomic clone of the yeast *CDC28* gene, was isolated by introduction of a *S.cerevisiae* genomic library made in λ YES (S.W.Ramer, S.J.Elledge and R.W.Davis, unpublished) into Y10 and selecting for complementation of the *cdc28-4* allele. Restriction analysis of the clone recovered from this strain agreed with the published restriction map (Reed *et al.*, 1982). Y62 (*CDC28/cdc28::TRP1*) was constructed by disruption of one of the two *CDC28* genes of CRY3 (*MATa/MATα ura3-1/ura3-1, trp1-1/trp1-1, his3-11,15/his2-11,15, ade2-1/ade2-1, leu2-3,112/leu2-3,112, can1-100/can1-100*) (a gift from B.Fuller, Stanford), with a fragment of DNA from pSE1040, containing the *TRP1* gene inserted into the *Sty*I site of the *CDC28* gene. Disruptions were confirmed by Southern analysis. Y82 (*MATa ade2-1 cdc28-4 ura3 leu2-3,112*) was derived from the sporulation of a diploid generated from a cross between CRY2 (*MATα his3-11,15 trp1-1 can1-100 ade2-1 ura3-1 leu2-3,112*) (B.Fuller, Stanford) and 4018-28-1 (*MATa cdc28-4 leu2-3,112 ura3*) (M.Cai, Seattle). The 5' PCR primer used to generate the 5' end of CDK2 cloned in pSE1037 was 5'-CCTCGAGCCACCATGGAGACCTT-CCAA-3'. The 3' primer was derived from vector sequences and a *Bam*HI site in the 3' untranslated region was used as the 3' restriction site for cloning purposes.

Media, enzymes, assays and genetic methods

For drug selection, LB plates were supplemented with ampicillin (50 mg/ml). Restriction endonucleases, *E.coli* DNA polymerase I large fragment, T4 polynucleotide kinase, T4 DNA polymerase and T4 DNA ligase were purchased from New England Biolab. Deoxyribonucleotides and ATP were purchased from P-L Biochemicals. Drugs were purchased from Sigma.

cDNA library construction

Poly(A)$^+$ mRNA (5 μg) prepared from Epstein–Barr virus transformed human peripheral lymphocytes (a gift from D.Denny, Stratford) was used for production of cDNA as previously described. cDNA, ~4 μg, was resuspended in 7 μl of TE and was ligated to 2 μg of an equal mixture of two different adaptors in a total volume of 10 μl at 4°C overnight. The sequence of the kinased adaptors was: adaptor 1, top strand 5'-CGAGATTTACC-3' and bottom strand 5'-GGTAAATC-3'; adaptor 2, top strand 5'-CGAGAGTTCAC-3' and bottom strand 5'-GTGAACTC-3'. cDNA was spermine precipitated and gel purified on a 1% low melting point agarose gel (Seakem) selecting cDNA >700 bp. 0.1 μg of cDNA was ligated to 2 μg of *Xho*I cleaved λ YES-R plasmid DNA that was modified by filling-in in the presence of only dTTP as described in a volume of 4 μl at 4°C overnight and packaged using one Gigapack Gold packaging extract (Stratagene, La Jolla, CA). 2×10^7 independent recombinants were obtained. This phage library was amplified on LE392 containing pMC9, a clone of *lacI* on pBR322, to reduce expression of the *lac* promoter. 2×10^8 phage were converted into plasmid by infection into BNN132 (Elledge *et al.*, 1991) and selecting for ampicillin resistant colonies. These colonies were pooled, grown to saturation in 6 l of LB supplemented with 50 μg/ml ampicillin, and plasmid DNA was prepared by CsCl density centrifugation. Transformation into yeast was performed as described previously (Elledge *et al.*, 1991).

DNA sequencing and Northern analysis

Plasmid DNAs were sequenced by the method of Sanger *et al.* (1977). Single-stranded plasmid DNAs were prepared by the method of Zagursky and Berman (1984) using R408 as a helper phage (Russel *et al.*, 1986). Hep 3B2 RNA was a gift from D.Wilson (Baylor). Northern blot analysis was performed as described previously (Elledge and Davis, 1990).

Acknowledgements

We thank M.Kuroda and W.Harper for critical comments on the manuscript. We would also like to thank D.Denny, B.Fuller, Z.Chen, M.Cai, C.Mann and D.Wilson for gifts of RNAs and yeast strains and many helpful discussions, and J.Tsuji and K.Matsumoto for sharing unpublished results. This work was supported by grant NIGMS-1R01GM44664-01 to S.J.E.

References

Aden,D.P., Foge,A., Plotkin,S., Damjanow,I. and Knowles,B.B. (1979) *Nature*, **282**, 615–616.
Blow,J.J. and Nurse,J.J. (1990) *Cell*, **62**, 855–862.
Boeke,J.D., Truehart,J., Natsoulis,G. and Fink,G.R. (1988) *Methods Enzymol.*, **154**, 164–175.
Draetta,G. and Beach,D. (1988) *Cell*, **54**, 17–26.
Elledge,S.J. and Davis,R.W. (1987) *Mol. Cell. Biol.*, **7**, 2783–2793.
Elledge,S.J. and Davis,R.W. (1990) *Genes Dev.*, **4**, 740–751.
Elledge,S.J., Mulligan,J.T., Ramer,S.W., Spottswood,M. and Davis,R.W. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1731–1735.
Evans,T. Rosenthal,E.T., Youngblom,J., Distel,D. and Hunt,T. (1983) *Cell*, **33**, 389–396.
Hadwiger,J.A., Wittenberg,C. and Mendenhall,M.D. (1989a) *Proc. Natl. Acad. Sci. USA*, **86**, 6255–6259.
Hadwiger,J.A., Wittenberg,C., Mendenhall,M.D. and Reed,S.I. (1989b) *Mol. Cell. Biol.*, **9**, 2034–2041.
Hartwell,L.H., Culotti,J., Pringle,J.R. and Reid,B.J. (1974) *Science*, **183**, 46–51.
Hayles,J., Beach,D., Durkatz,B. and Nurse,P. (1986) *Mol. Gen. Genet.*, **202**, 291–293.
Hindley,J. and Phear,G.A. (1984) *Gene*, **31**, 129–134.
Jimenez,J., Alphey,L., Nurse,P. and Glover,D.M. (1990) *EMBO J.*, **9**, 3565–3571.
Krek,W. and Nigg,W.A. (1989) *EMBO J.*, **8**, 3071–3078.
Lee,M.G. and Nurse,P. (1987) *Nature*, **327**, 31–35.
Lehner,C.F. and O'Farrell,P.H. (1990) *EMBO J.*, **9**, 3573–3581.
Lorincz,A.T. and Reed,S.I. (1984) *Nature*, **307**, 183–185.
McGowan,C.H., Russell,P. and Reed,S.I. (1990) *Mol. Cell. Biol.*, **10**, 3847–3851.
Nash,R., Tokiwa,G., Anand,S., Erikson,K. and Futcher,B. (1988) *EMBO J.*, **7**, 4335–4346.
Nurse,P. and Bissett,Y. (1981) *Nature*, **292**, 448–460.
Paris,J., Le Guellec,R.L., Couturier,A., Le Guellec,K., Omilli,F., Camonis,J., MacNeill,S. and Phillipe,M. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1039–1043.
Pines,J. and Hunter,T. (1990a) *Nature*, **346**, 760–763.
Pines,J. and Hunter,T. (1990b) *New Biologist*, **2**, 389–401.
Pondaven,P., Meijer,L. and Beach,D. (1990) *Genes Dev.*, **4**, 9–17.
Reed,S.I. (1990) *Genetics*, **95**, 561–577.
Reed,S.I., Ferguson,J. and Groppe,J.C. (1982) *Mol. Cell. Biol.*, **2**, 412–425.
Riabowol,K., Draetta,G., Brizuela,L., Vandre,D. and Beach,D. (1989) *Cell*, **57**, 393–401.
Russel,M., Kidd,S. and Kelly,M.R. (1986) *Gene*, **45**, 333–338.
Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) *Science*, **239**, 487–491.
Sanger,F., Nicklen,F. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Spurr,N.K., Gough,A.C. and Lee,M.G. (1990) *J. DNA Sequencing and Mapping*, **1**, 49–54.
Standart,N., Minshull,J., Pines,J. and Hunt,T. (1987) *Dev. Biol.*, **124**, 248–254.
Swenson,K.I., Farrell,K.M. and Ruderman,J.V. (1986) *Cell*, **47**, 861–870.
Th'ng,J.P.H., Wright,P.S., Hamaguchi,J., Lee,M.G., Norbury,C.J., Nurse,P. and Bradbury,E.M. (1990) *Cell*, **63**, 313–324.
Westendorf,J.M., Swenson,K.I. and Ruderman,J.V. (1989) *J. Cell Biol.*, **108**, 1431–1444.
Wittenberg,C. and Reed,S.I. (1989) *Mol. Cell. Biol.*, **9**, 4064–4068.
Wittenberg,C., Sugimoto,K. and Reed,S.I. (1990) *Cell*, **62**, 225–237.
Yannisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Zagurski,R.J. and Berman,M.L. (1984) *Gene*, **27**, 183–191.

Received on March 7, 1991; revised on May 24, 1991