Molecular Characterization of Cell Cycle Gene CDC7 from Saccharomyces cerevisiae

MARK PATTERSON,1 ROBERT A. SCLAFLANI,2†† WALTON L. FANGMAN,2 AND JOHN ROSAMOND1

Department of Biochemistry, University of Manchester, Manchester M13 9PT, United Kingdom,1 and Department of Genetics, University of Washington, Seattle, Washington 981952

Received 5 November 1985/Accepted 13 January 1986

The product of the CDC7 gene of Saccharomyces cerevisiae appears to have multiple roles in cellular physiology. It is required for the initiation of mitotic DNA synthesis. While it is not required for the initiation of meiotic DNA replication, it is necessary for genetic recombination during meiosis and for the formation of ascospores. It has also been implicated in an error-prone DNA repair pathway. Plasmids capable of complementing temperature-sensitive cdc7 mutations were isolated from libraries of yeast genomic DNA in the multicopy plasmid vectors YRp7 and Yep24. The complementing activity was localized within a 3.0-kilobase genomic DNA fragment. Genetic studies that included integration of the genomic insert at or near the CDC7 locus and marker rescue of four cdc7 alleles proved that the cloned fragment contains the yeast chromosomal CDC7 gene. The RNA transcript of CDC7 is about 1,700 nucleotides. Analysis of the nucleotide sequence of a 2.1-kilobase region of the cloned fragment revealed the presence of an open reading frame of 1,521 nucleotides that is presumed to encode the CDC7 protein. Depending on which of two possible ATG codons initiates translation, the calculated size of the CDC7 protein is 58.2 or 56 kilodaltons. Comparison of the predicted amino acid sequence of the CDC7 gene product with other known protein sequences suggests that CDC7 encodes a protein kinase.

Use of the budding yeast Saccharomyces cerevisiae as a model for studies on the eucaryotic cell cycle relies heavily on temperature-sensitive mutations in cell division cycle (CDC) genes (32, 35). Characterization of these mutants has led to the formulation of a model in which progression through the cell cycle is determined by a set of interrelated pathways, each organized as a dependent sequence of events requiring the action of specific gene products. One such pathway, operating late in the G1 phase of the cell cycle, requires the function of the CDC7 gene product (12). Cells carrying a thermosensitive lesion in the CDC7 gene arrest at the restrictive temperature as budded cells with separated spindle-pole bodies but without an elongated spindle apparatus and without initiating DNA synthesis (5, 12). Upon return to permissive conditions cdc7 cells are able to enter the S phase and subsequently complete a round of DNA synthesis without further protein synthesis (14).

In contrast to the requirement for CDC7 function to initiate mitotic DNA synthesis, meiotic DNA replication occurs normally in cdc7 homozygous diploids under the restrictive condition (43). However, these diploids fail to form a synaptonemal complex, to show commitment to genetic recombination, or to form ascospores (40). Thus, although cdc7 strains are defective in both mitotic and meiotic cell cycles, the lesion appears to affect each pathway in a quite distinct manner. In addition to having roles in the mitotic and meiotic pathways, the CDC7 gene product has been implicated in DNA repair as a member of the RAD6 epistasis group, since strains carrying a cdc7 mutation show almost no mutagenic repair in response to a variety of damaging agents (31).

To elucidate the role of the CDC7 gene product in the various cellular functions in which it is implicated and to determine whether differential expression of the CDC7 gene is involved with its cell cycle functions, we and others (24) have begun a molecular analysis of the CDC7 gene. In this paper we describe the cloning of the CDC7 gene, the characterization of its transcriptional product, the nucleotide sequence of the gene, and the regions of homology between the predicted protein products of the CDC7 and CDC28 genes.

MATERIALS AND METHODS

Strains and media. Escherichia coli HB101 (F− thi leu pro hisD hsdM recA end1) and HW87 [F− Δ(araD139-leu) lacX74 galK hsdR rpsL srb recA] were used as hosts for the routine maintenance and propagation of plasmids. Bacterial cultures were grown in L broth or supplemented M9 medium (26); when necessary, ampicillin was added to media to a final concentration of 50 μg/ml.

Yeast cells were grown in either yeast extract-peptone-glucose (YPD) or supplemented synthetic minimal medium (42). S. cerevisiae strains used in this work were SB155 (MATa trpl cdc7-1), SB158 (MATa trpl ura3-52 cdc7-1), S288C (MATα), 136 (MATα trpl ura3-52 leu2-3,112 cdc7-2), 142 (MATα trpl ura3-52 leu2-3,112 cdc7-4), 208 (MATα ura3-52 leu2-3,112 cdc7-1), and 209 (MATα ura3-52 leu2-3,112 cdc7-3). Strains containing the cdc7 alleles were constructed by standard genetic procedures (42) from the original cdc7 isolates obtained from L. H. Hartwell (13).

Preparation of DNA. Plasmid DNA was extracted from cultures of E. coli on an analytical scale by alkaline lysis (2) and purified from larger cultures by CsCl-ethidium bromide
density gradient centrifugation after detergent lysis (16). DNA was prepared from rapid lysates of yeast transformants essentially as described by Naumovski and Friedberg (30), while genomic DNA was prepared from *S. cerevisiae* S288C by the method of Cryer et al. (8). Single-stranded viral DNA was prepared from recombinant M13mp8 and M13mp9 bacteriophage by phenol extraction of polyethylene glycol-precipitated phage particles (25).

**Characterization of DNA.** Restriction enzymes, T4 DNA ligase, DNA polymerase, and Klenow fragment were purchased from New England Biolabs, Inc. (Keene, N.H.) or Bethesda Research Laboratories, Inc. (Gaithersburg, Md.) and were used according to the recommendations of the manufacturers. DNA was nick translated as described by Rigby et al. (37) with [α-32P]dATP supplied by Amersham International and New England Nuclear Corp. (Boston, Mass.).

**Yeast genomic libraries.** Pools of recombinant plasmids containing quasi-random fragments of yeast genomic DNA (strain S288C) were prepared essentially as described previously, using the vector YRp7 (29, 32, 45). The Carlson YEp24 genomic library (6) was obtained from D. Koshland and D. Botstein.

**DNA sequencing.** Nucleotide sequences were determined by the dideoxy chain-termination method (38), for which DNA fragments to be sequenced were cloned into M13mp8 or M13mp9 (25). Reaction products were resolved by electrophoresis through 6% acrylamide gels under denaturing conditions and detected by autoradiography overnight at room temperature.

**Genetic techniques.** Transformation of *E. coli* by plasmid DNA was by the method of Warren and Sherratt (48). *S. cerevisiae* strains were transformed after spheroplasting (42). Construction of diploids, sporulation, and dissection of yeast spore tetrads were performed by standard genetic techniques (42).

**Marker rescue analysis.** Plasmid pRS4 was produced from pRS3 by digestion with BamHI and subsequent recircularization and self-ligation. Plasmid pRS5 was produced by subcloning the 5.2-kilobase (kb) BamHI fragment from pRS3 into the BamHI site of vector YEp13 (4). All cdc7 mutant strains were complemented by plasmid pRS3 (7.4 kb) but not by plasmids pRS4 (2.2 kb) and pRS5 (5.2 kb). For the marker rescue analysis, Ura+ (or Leu+) transformant colonies selected at 23°C were picked, and the cells were plated at 23 and 36°C on fully supplemented medium (YPD). The frequency of reversion was calculated as the number of colonies produced at the restrictive temperature (36°C) divided by the total number of colonies which harbored the plasmid at the permissive temperature (23°C). Usually about 80% of the cells contained the plasmid under these conditions. Values in Table 1 are averages of five trials on each of two independent transformants.

**Preparation of RNA from *S. cerevisiae*.** Total RNA was prepared from 50-ml cultures of exponentially growing yeast cells at a density of 2 × 107 cells per ml as follows. Cells were harvested by YPD, washed once in ice-cold water, and suspended in 3 ml of ice-cold breaking buffer (0.1 M Tris [pH 7.5], 0.1 M LiCl, 0.1 mM EDTA, 0.5 mg of heparin per ml). An equal volume of glass beads was added, and the cells were broken by vortexing for four periods of 30 s with cooling on ice in between. Sodium dodecyl sulfate was then added to a final concentration of 0.5%, and the aqueous phase was extracted with an equal volume of phenol-

**TABLE 1. Marker rescue analysis of four cdc7 mutant alleles**

<table>
<thead>
<tr>
<th>cdc7 allele</th>
<th>Control reversion (x10^(-5))</th>
<th>pRS5 reversion (x10^(-5))</th>
<th>Relative increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-1</td>
<td>0.23</td>
<td>1.6</td>
<td>7.0</td>
</tr>
<tr>
<td>7-2</td>
<td>0.58</td>
<td>2.8</td>
<td>4.8</td>
</tr>
<tr>
<td>7-3</td>
<td>1.50</td>
<td>180</td>
<td>120</td>
</tr>
<tr>
<td>7-4</td>
<td>1.50</td>
<td>150</td>
<td>100</td>
</tr>
</tbody>
</table>

a Strains were transformed with control plasmid pRS3 (*CDC7*) or with plasmid pRS5 which contains the 5.2-kb DNA fragment located at the left of the BamHI site in the *CDC7*-complementing region of pRS3 (Fig. 1; Materials and Methods).

b Reversion frequencies were not significantly different for strains without plasmid and the same strains harboring the vector plasmid or pRS4.

c The increase in the frequency of reversion produced by strains harboring plasmid pRS5.

| 80% of the cells contained the plasmid under these conditions. Values in Table 1 are averages of five trials on each of two independent transformants. **Preparation of RNA from *S. cerevisiae*.** Total RNA was prepared from 50-ml cultures of exponentially growing yeast cells at a density of 2 × 107 cells per ml as follows. Cells were harvested by YPD, washed once in ice-cold water, and suspended in 3 ml of ice-cold breaking buffer (0.1 M Tris [pH 7.5], 0.1 M LiCl, 0.1 mM EDTA, 0.5 mg of heparin per ml). An equal volume of glass beads was added, and the cells were broken by vortexing for four periods of 30 s with cooling on ice in between. Sodium dodecyl sulfate was then added to a final concentration of 0.5%, and the aqueous phase was extracted with an equal volume of phenol-

**RESULTS**

**Isolation of DNA fragments able to complement the cdc7 mutation.** Genomic DNA fragments capable of complementing the temperature-sensitive cdc7 mutation were isolated from a library of random Sau3A fragments of *S. cerevisiae* DNA in the vector YEp24. Plasmid YEp24 consists of pBR322 carrying the yeast *URA3* gene and the 2μm plasmid replication origin (3); the recombinant plasmids of the library replicate autonomously in yeasts and express the *URA3* gene. A *ura3 cdc7* strain of *S. cerevisiae* (strain 136) was transformed with the library, and Ura+ transformants were selected by allowing spheroplasts to regenerate in agar medium lacking uracil at 23°C for 5 days. The required Ura+ TSM*+* transformants were identified by replica plating Ura+ transformants to fresh agar medium lacking uracil and incubating at 36°C. Two Ura+ TSM*+* transformants of *S. cerevisiae* were obtained in this way, and total nucleic acid extracted from each of these was used to transform *E. coli* HB101 to ampicillin resistance. Isolation and characterization of the transforming plasmids indicated that they carried nonidentical but overlapping genomic fragments. These plasmids were designated pRS3 and pRS7 (Fig. 1).

Genomic fragments capable of complementing cdc7 were also isolated independently from a Sau3A genomic library in the vector YRP7. This vector carries the yeast *TRP1* gene and a putative chromosomal replication origin that permits autonomous replication in yeasts (45). In this case, a trp1 cdc7 strain (strain SB155) was transformed with the library, and transformants that grew at 36°C on agar medium lacking tryptophan were selected directly. Total DNA was extracted from two such yeast transformants and was used to trans-
form *E. coli* HB101 to ampicillin resistance. The plasmids recovered in this manner were designated pMP101 and pMP201 (Fig. 1).

Restriction enzyme mapping showed that the genomic fragments in each of the complementing plasmids were different. pRS3 and pRS7 contained genomic fragments of 8.2 and 10.9 kb, respectively, while pMP101 and pMP201 carried inserts of 3.0 and 8.1 kb, respectively. There was significant overlap between the four cloned genomic fragments, and in particular, the insert of pMP101 was found to be contained within the three other plasmids which complemented *cdc7* (Fig. 1). Southern analysis of chromosomal DNA showed that the distribution of sites in the cloned DNA fragments of pRS3 and pRS7 was the same as at the homologous region of the genome of *S. cerevisiae* (data not shown). Figure 1 shows a composite map of restriction sites within this region of the genome.

Closed fragments contain the authentic *CDC7* gene. Although the isolation of homologous genomic fragments capable of complementing *cdc7* from two independently constructed genomic libraries suggests that the cloned fragments carry the *CDC7* gene itself, the observation that the effects of many conditional lethal mutations in yeasts can be alleviated by extragenic suppression (20) made it necessary to demonstrate further that the plasmids contained the authentic *CDC7* gene within the cloned genomic fragment. The genomic fragment from pMP101 was first subcloned into the vector YIp5, which carries the yeast *URA3* gene in PBR322. Since this recombinant plasmid, designated pMP104 (Fig. 1), is unable to replicate autonomously in yeasts (39), stable yeast transformants arise only if the plasmid integrates into a chromosome by homologous recombination (15). Thus, plasmid pMP104 was used to transform a *ura3* *cdc7* strain (strain SB158), and *URA* + *TSM* + transformants were selected directly. Since the *ura3*-52 mutation appears to preclude recombination between this locus and the vector *URA3* + sequences (39), integration would be likely to occur at the chromosomal site homologous to the genomic fragment of pMP104.

A standard genetic cross was then performed to determine whether pMP104 had indeed integrated at the *CDC7* locus, which is tightly linked to both the *TRP1* locus (3.8 centimorgans) and the centromere of chromosome IV (27, 28). One integrant (*MATa his ade trp1 ura3-52 cdc7-1::URA3 + TSM +*) was crossed with SB107 (*MATa leu2-3 ura3-1*); tetrads were dissected and scored for TSM, URA, and TRP phenotypes. In 25 of 25 tetrads, TSM + segregated 4+:0−. Twenty-four tetrads showed parental ditype, and one showed tetratype segregation for URA and TRP, indicating that the integrated sequences are about 2 centimorgans from *TRP1* (33). We conclude that plasmid pMP104 had integrated at or near the *CDC7* locus. Therefore, plasmids pMP101 and pMP104 carry a genomic insert which is most likely the *CDC7* gene.

**Location of *CDC7* gene within cloned DNA fragments.** The observation that the genomic fragment of pMP101 is common to the cloned fragments of pRS3, pRS7, and pMP201 suggests that the complete *CDC7* gene is contained entirely within this 3.0-kb segment of DNA. To localize the functional *CDC7* gene more precisely, we prepared two subclones of the genomic DNA insert of pMP101. In the first of these, the 650-base-pair (bp) fragment between the new *BamHI* site created at the vector-insert junction and the natural *BamHI* site was removed by digesting pMP101 to completion with *BamHI* and religating after dilution. The resulting plasmid, pMP102 (Fig. 1), transformed *S. cerevisiae* to *TRP* + at high frequency at 23°C but was unable to complement the *cdc7* mutation, suggesting that the *CDC7* gene extends to the right of the authentic *BamHI* site of pMP101 as drawn in Fig. 1.

Plasmid pMP106 was constructed by digesting pMP101 to completion with *ClaI* and *SacI*, removing single strands with

---

**FIG. 1.** Restriction map of *cdc7*-complementing clones and subcloned DNA fragments. The upper line represents a composite of the data for individual clones. Abbreviations: +, able to complement *cdc7*; −, unable to complement *cdc7*; (B), new *BamHI* sites created at the vector-insert junctions; B, *BamHI* site; C, *ClaI* site; G, *BglII* site; H, *HindIII* site; M, *MluI* site; P, *PstI* site; R, *EcoRI* site; S, *SacI* site; Sp, *SphI* site; V, *EcoRV* site.
the BamHI site (Fig. 1). Although the four cdc7 alleles were isolated as independent clones (12), it is possible that they represent the same mutation. However, it is evident from the data presented in Table I that at least cdc7-1 and cdc7-2 must be different from cdc7-3 and cdc7-4 because the frequency of recombination with pRS5 was very different (about 20-fold). Further deletion mapping revealed that at least three of the alleles, cdc7-1, -2, and -3, are located between the BamHI and EcoRV sites on a 0.77-kb DNA fragment (data not shown).

Identification of CDC7 RNA transcript. RNA was prepared from wild-type S. cerevisiae and was denatured with treatment by glyoxal (46). Three equal samples of RNA were fractionated by electrophoresis through 1.5% agarose gels, and the RNA was transferred to nitrocellulose as described previously (46). The nitrocellulose filter was divided into three separate strips, and each was hybridized with a nick-translated DNA probe as indicated in Fig. 2.

Four different RNA species were identified as being homologous to regions of the cloned DNA around the CDC7 gene. The 1.3-kb BamHI-ClaI fragment hybridized to a single RNA species of 1,700 nucleotides (Fig. 2, lane 2). In other experiments, the 0.77-kb BamHI-EcoRV fragment which could rescue three cdc7 alleles (see above) also hybridized to a 1,700-nucleotide RNA which is poly(A)+ (data not shown). From these data we conclude that the 1,700-nucleotide RNA is the transcript of the CDC7 gene, which is consistent with a previous report (24). The 650-bp DNA fragment between the BamHI site and the new BamHI site at the vector-insert boundary also hybridized to a 1,700-nucleotide RNA species (Fig. 2, lane 1), which is consistent with the results of a subcloning experiment that showed that at least part of this region was necessary for CDC7 expression. The 650-bp DNA probe also hybridized to a 1,300-nucleotide RNA that we presume to be the transcript of a gene adjacent to CDC7. The 1.5-kb ClaI-ClaI fragment hybridized to two RNA transcripts of about 1,000 and 900 nucleotides (Fig. 2, lane 3). We did not detect hybridization between this fragment and the 1,700-nucleotide RNA, which supports the view that regions to the left of the ClaI site in pMP101 are not essential for complementation of the cdc7 mutation. However, we cannot exclude the presence of a short region of homology between the 1.5-kb ClaI-ClaI fragment and the putative CDC7 transcript.

Nucleotide sequence of the CDC7 gene. The nucleotide sequence of the genomic fragment in the region of the CDC7 gene was determined initially from the new BamHI site at one vector-insert junction in pMP101 to the ClaI site and subsequently to a point 110 nucleotides beyond ClaI, using the strategy shown in Fig. 3A. When possible, restriction subfragments of the genomic clone were sequenced directly. Complete sequencing, however, required the analysis of randomly isolated subclones of RsaI and Sau3A fragments.

The location of all potential termination codons in this region is shown in Fig. 3B, from which it can be seen that only one of the possible reading frames contains a long stretch (507 triplets) uninterrupted by stop codons. The size of this long reading frame (1,521 bases) is compatible with the size of the CDC7 mRNA transcript (~1,700 bases; Fig. 2). Moreover, the reading frame contains both the BamHI and SacI sites, consistent with the inability of clones terminating at these sites to complement the cdc7 mutation (Fig. 1). This reading frame also contains the 0.77-kb BamHI-EcoRV fragment which can rescue three cdc7 alleles (see above). From these data we conclude that this open reading frame encodes the CDC7 gene product.
Figure 4 shows the nucleotide sequence of a 2.1-kb region around the long open reading frame, together with a predicted amino acid sequence for the translational product. The nucleotide A of the first in-phase ATG codon within the open reading frame is numbered 1. However, it is not clear that this ATG codon is the translational initiation site, since there is a second in-frame ATG codon at nucleotides 55 to 57 that could act as the initiator (Fig. 4). Depending on which of these codons is used for initiation, the calculated molecular weight of the CDC7 protein is either 58,250 or 56,000.

The nucleotide sequence downstream from the TAG termination codon (1522 to 1524) contains several short nucleotide sequences found in the 3′-untranslated regions of other S. cerevisiae genes, notably the sequences TAGT and TTT, which occur repetitively between nucleotides 1544 and 1678. The TAGT at nucleotides 1544 to 1547 is part of the sequence TAGTCT that occurs repetitively downstream of the CDC8 gene (1) and may be associated with transcription termination and polyadenylation.

**Homology between CDC7 gene product and protein kinases.**

The predicted amino acid sequence of the CDC7 gene product was compared with other known and predicted protein sequences. This search revealed a statistically significant homology between the CDC7 protein and the protein product of the CDC28 gene, which has been shown to be a protein kinase (22, 36). The homology is not randomly distributed throughout the proteins though, but is confined largely to four domains comprising residues 40 to 52, 73 to 78, 155 to 186, and 275 to 308 of the CDC7 sequence (Fig. 5). These domains correspond to two functionally important regions within the CDC28 and other protein kinases, one around an ATP-binding site and the other surrounding a putative phosphorylation receptor site (11, 17, 22). Both these sites are thought to be essential for protein kinase activity, and their amino acid sequences are highly conserved in a number of known and putative kinases including bovine cyclic AMP-dependent protein kinase and the src family of oncogene kinases (11, 17, 22).

The majority of the consensus sequence information for both sites found in protein kinases was present in the CDC7 gene product (Fig. 6) and constituted the region of maximum homology with CDC28. Thus, codons 40 to 52 and 73 to 78 approximate the ATP-binding site, while codons 155 to 186 and 275 to 308 encompass the phosphorylation receptor site. These homologies suggest that CDC7 is a protein kinase. However, while the CDC7 protein contains the consensus
sequence information of the protein kinase functional domains, the organization of that sequence differs significantly from the consensus. In particular, the regions within each site that exhibit length heterogeneity are larger in CDC7 by some 10 amino acids in the ATP-binding site and by about 80 amino acids at the phosphorylation receptor site.

**DISCUSSION**

Four plasmids capable of complementing the cdc7 mutation were isolated from libraries of *S. cerevisiae* genomic DNA. One of these plasmids (pMP101) carried an insert of 3.0 kb that was present within the cloned fragments of the
other three plasmids. The cdc7-complementing activity was localized within this fragment by subcloning and shown to correspond to an open reading frame of 1,521 bp that is transcribed to produce an mRNA of 1,700 nucleotides.

The genomic fragment carrying this reading frame was subcloned into the integrative vector YIp5, and the resulting plasmid (pMP104) was used to transform an S. cerevisiae cdc7 strain such that the plasmid integrated into the yeast genome at the site homologous to the cloned insert. Tetrad analysis showed that this site maps at or near the known locus of the CDC7 gene, near the TRP1 gene on chromosome IV. Furthermore, by employing plasmid-chromosome recombination, the mutant sites in all of four cdc7 alleles (cdc7-1, -2, -3, and -4) were shown to be located in the regions corresponding to specific cloned DNA fragments (Fig. 1; Table 1). Therefore, the cloned fragments contain the authentic CDC7 gene and not an extragenic suppressor.

It is interesting to note that the recombinant plasmid pMP104 transformed S. cerevisiae at the very low frequency (1 to 5 transformants per μg of DNA) characteristic of yeast vectors lacking the sequences necessary for autonomous replication (ARS elements [7, 45]), implying that neither YIp5 nor the cloned fragment carrying CDC7 contain a functional ARS element. However, the nucleotide sequence of the cloned fragment reveals a sequence that is closely related to the consensus core sequence, —TTTATPuTT T—, found within identified ARS elements (4, 18, 44, 47). The sequence —ATTTGTATTT—, complementary to nucleotides 1795 to 1805 in the 3′-flanking region of the CDC7 gene, differs by only one nucleotide from the ARS core consensus. We presume that either the change from A to G at nucleotide 5 of this sequence is sufficient to prevent ARS activity or that adjacent sequences necessary for autonomous replication are lacking (7, 18).

In addition to the approximate consensus ARS element, the 3′-flanking region of the CDC7 gene contains sequences found in similar regions of other yeast genes, notably the consensus polyadenylation sequence AATAAA (10) at nucleotides 1619 to 1624 (Fig. 4), 95 nucleotides downstream from the TAG termination codon. It is part of a larger
sequence, TAG—(A-T rich)₆—TTT, occurring between nucleotides 1605 and 1629 that strongly resembles a sequence Zaret and Sherman (50) have proposed to be required for efficient transcription termination and polyadenylation.

In contrast to the 3′-flanking region, there is no evidence for any of the consensus signal sequences (TATAAA and PuCACACA) that occur in the 5′-flanking regions of many yeast genes (49). This failure may be a consequence of the short amount of DNA sequenced 5′ to the open reading frame; while the consensus TATA box signal is located about 35 bp upstream of the transcription start site in most eucaryotes, it may be as far as 220 bp upstream in yeasts (41).

The open reading frame itself could encode two proteins of molecular weight 58,250 or 56,000 depending on which of two ATG codons is used for the initiation of translation. Both ATG codons, at nucleotides 1 to 3 and 55 to 57, are located within sequences that approximate the PuXXATGpuXT that is found in the translational start site of many S. cerevisiae genes (19). It may be significant that a plasmid construct derived from pMP101 in which the genomic insert terminates at the Clal site (nucleotides 22 to 27) retains the ability to transform S. cerevisiae to TRP⁺ and TSM⁺ at high frequency (M. N. Patterson, unpublished data). Since this plasmid lacks genomic sequences upstream of nucleotide 22 (Fig. 4), it seems likely that in this case the ATG codon at nucleotides 55 to 57 is used for translational initiation. This observation would be consistent with the Northern hybridization studies (Fig. 2) that suggested that the CDC7 mRNA did not extend beyond the Clal site. However, this interpretation would require that all the regulatory and promoter sequences for CDC7 expression be located between nucleotides 22 and 54 of the sequence shown in Fig. 4, which seems unlikely. Another possibility is that in this construct CDC7 expression is dependent on adjacent vector sequences, such that a truncated but functional protein is being produced. Definitive conclusions on this point require accurate mapping of the 5′ terminus of the CDC7 transcript.

The CDC7 protein sequence has regions of homology with the CDC28 and oncogene protein kinases (Fig. 6) (11, 17, 22). However, CDC7 differs from all known protein kinases by virtue of a large region of heterology within the phosphorylation receptor domain. One explanation for this could be the presence of introns within the CDC7 gene. However, only one RNA species has been observed in Northern hybridization studies, and CDC7 lacks any of the consensus exon-intron junction and 3′ splice signal sequences (21, 34), implying that the extra amino acid sequences predicted within the phosphorylation receptor domain are indeed present in the CDC7 gene product. This being the case, it is less obvious that these domains are associated within a protein kinase activity of the CDC7 protein. The additional heterogeneity might modify the phosphorylation receptor domain such that the protein performs functions unrelated to protein phosphorylation. Alternatively, the CDC7 protein may have protein kinase activity that is modified or regulated in a specific manner as a result of the heterogeneity.

The demonstration that CDC28 encodes a protein kinase (36) suggests that commitment to the mitotic cell cycle is accomplished in part by the activation of target proteins by phosphorylation. The finding that the CDC7 protein may also be a protein kinase suggests that the initiation of mitotic DNA synthesis also requires the phosphorylation of certain specific proteins. Moreover, it implies that the nuclear division pathway operating in the late G1 phase is associated with events that resemble a cascade of protein phosphorylation that result ultimately in the transition from G1 to the S phase. The identification of other components of the cascade and the characterization of the targets for phosphorylation by CDC7 and CDC28 protein kinases would help test this idea.

ACKNOWLEDGMENTS

We thank Steve Field, Bruce Venning, Dave Moore, and Majeed Bahman for useful discussions, Breck Byers, Bonita Brewer, and Robert McC Carroll for comments on the manuscript, and Margaret Barber and Nancy Gamble for secretarial assistance. We are especially grateful to Douglas Wagner for helping with subcloning and early characterization of the CDC7 plasmids.

This work was supported by an SERC studentship (to M.P.), an SERC Project grant (to J.R.), a National Institutes of Health Research Service award (to R.A.S.), and a Public Health Service grant (to W.L.F.).

LITERATURE CITED