*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 12475–12479, December 1995
Genetics

# Orp1, a member of the Cdc18/Cdc6 family of S-phase regulators, is homologous to a component of the origin recognition complex

M. MUZI-FALCONI AND THOMAS J. KELLY

Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205

*Contributed by Thomas J. Kelly, September 11, 1995*

**ABSTRACT** ***cdc18*$^+$ of *Schizosaccharomyces pombe* is a periodically expressed gene that is required for entry into S phase and for the coordination of S phase with mitosis. *cdc18*$^+$ is related to the *Saccharomyces cerevisiae* gene *CDC6*, which has also been implicated in the control of DNA replication. We have identified a new *Sch. pombe* gene, *orp1*$^+$, that encodes an 80-kDa protein with amino acid sequence motifs conserved in the Cdc18 and Cdc6 proteins. Genetic analysis indicates that *orp1*$^+$ is essential for viability. Germinating spores lacking the *orp1*$^+$ gene are capable of undergoing one or more rounds of DNA replication but fail to progress further, arresting as long cells with a variety of deranged nuclear structures. Unlike *cdc18*$^+$, *orp1*$^+$ is expressed constitutively during the cell cycle. *cdc18*$^+$, *CDC6*, and *orp1*$^+$ belong to a family of related genes that also includes the gene *ORC1*, which encodes a subunit of the origin recognition complex (ORC) of *S. cerevisiae*. The products of this gene family share a 250-amino acid domain that is highly conserved in evolution and contains several characteristic motifs, including a consensus purine nucleotide-binding motif. Among the members of this gene family, *orp1*$^+$ is most closely related to *S. cerevisiae ORC1*. Thus, the protein encoded by *orp1*$^+$ may represent a component of an *Sch. pombe* ORC. The *orp1*$^+$ gene is also closely related to an uncharacterized putative human homologue. It is likely that the members of the *cdc18/CDC6* family play key roles in the regulation of DNA replication during the cell cycle of diverse species from archaebacteria to man.**

The initiation of DNA replication is tightly controlled during the eukaryotic cell cycle (1–7). Initiation takes place at multiple origins of replication distributed along the chromosomal DNA and is normally triggered when cells reach a characteristic size. A key feature of the initiation process is that origins are activated once and only once each cell cycle, ensuring that exactly two copies of each segment of chromosomal DNA are produced. Recent work has resulted in the identification of a protein complex in budding yeast that specifically recognizes origins of DNA replication (8). Several lines of evidence indicate that this origin recognition complex (ORC) is necessary for the initiation of DNA replication (8–11). However, DNA footprinting suggests that ORC may be bound to chromosomal origins not just in S phase, but throughout the cell cycle (12). Prior to S phase, the ORC footprint is extended by an additional region of protection (12). Thus, it is likely that additional factors contribute to triggering the onset of DNA replication at S phase.

In the fission yeast *Schizosaccharomyces pombe*, the transition from $G_1$ to S phase of the cell cycle requires the transient activation of a transcription factor containing the products of the *cdc10*$^+$ and *sct1*$^+$ genes (13). Activation occurs in late $G_1$ and is dependent upon the activity of the Cdc2 protein kinase, which is known to regulate both the $G_1$/S and $G_2$/M transitions (14). A key target of the Cdc10/Sct1 transcriptional activator is the *cdc18*$^+$ gene (1, 15). Expression of *cdc18*$^+$ from a heterologous promoter is sufficient to rescue the lethality of a conditional temperature-sensitive (ts) *cdc10*$^{ts}$ mutant. The *cdc18*$^+$ gene product, a 65-kDa protein, is essential for the $G_1$/S transition. Moreover, p65$^{cdc18}$ is a highly labile protein whose expression is confined to a narrow window at the $G_1$/S boundary (unpublished data). These properties are consistent with the hypothesis that Cdc18 may play an important role in regulating the initiation of DNA replication at S phase. The Cdc18 protein is homologous to the budding yeast Cdc6 protein, which may have a similar function (16).

In this report we describe the identification and characterization of *orp1*$^+$,* a *Sch. pombe* gene related to *cdc18*$^+$. The protein encoded by *orp1*$^+$ is even more closely related to the large subunit of the budding yeast ORC, suggesting that it may represent a component of a similar complex in *Sch. pombe*. Cells lacking *orp1*$^+$ are inviable. When spores carrying a disrupted version of *orp1*$^+$ are germinated, they undergo one or a few rounds of DNA replication but do not progress further in the cell cycle. Many of the arrested cells show the elongated phenotype characteristic of *cdc* mutants and have abnormal nuclear structures. Interestingly, the Cdc18, Cdc6, Orc1, and Orp1 proteins share a 250-amino acid domain that is highly conserved in evolution. The family of proteins containing this domain likely plays an important role in the regulation of DNA replication.

## MATERIALS AND METHODS

**Strains and Plasmids.** All *Sch. pombe* strains used are isogenic to the wild-type 972 $h^-$. Strain TK79 $h^+/h^-$ *cdc18*$^+$/*cdc18*::*ura4*$^+$ *leu1-32/leu1-32 ura4-D18/ura4-D18 ade6-M210/ade6-M216* was used as a control in the spore germination experiment. Cells were transformed by electroporation as described (15). For temperature-shift experiments, cells were grown at the permissive temperature of 25°C and shifted to 35.5°C for the indicated times. Wild-type strains were grown at 30°C. Media were as described (17).

**Degenerate PCR.** *Sch. pombe* genomic DNA was prepared as described (17). Reaction mixtures (100 μl) contained 10 mM Tris·HCl (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl, 1 unit of *Taq* DNA polymerase (BRL), and 500 pmol of the following pair of degenerate primers: 5′-CCGGAATTCTAYRTNASNGGNSCICCIGGIIGIGGIAARAC-3′ and 5′-CCGGAATTCANNTBNARNGCICGICKIGYITCICC-3′, where T or C = Y; A or G = R; C or G = S; G or T = K; A, C, G, or T = N; and G, T, or C = B. PCR was carried out in a Robocycler Gradient 40 (Stratagene) under the following conditions: 1 min at 94°C, 1.5 min at 42–56°C gradient, 1 min at 72°C for 30 cycles, terminated with a final 5 min at 72°C. The PCR product was isolated from an agarose gel by using the QIAEX extraction kit (Qiagen, Chatsworth, CA) according to the manufac-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: ORC, origin recognition complex.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. U38522).

turer's instructions and was subcloned into the *Eco*RI site of pBluescript II SK(+), generating the plasmid pMF93.

**Isolation of the *orp1*$^+$ Gene.** The degenerate PCR reaction product was labeled with $^{32}$P by using the DECAPrimeII kit (Ambion, Austin, TX) and was used to probe an ordered *Sch. pombe* cosmid library as described (18). The *orp1*$^+$ gene was localized to cosmid 2G5c. Additional clones were isolated from *Sch. pombe* genomic and cDNA libraries that have been described elsewhere (15). Sequence assembly and similarity analyses were carried out by using SEQUENCHER (Gene Codes Corp., Ann Arbor, MI) and the Genetics Computer Group (GCG) package, respectively.

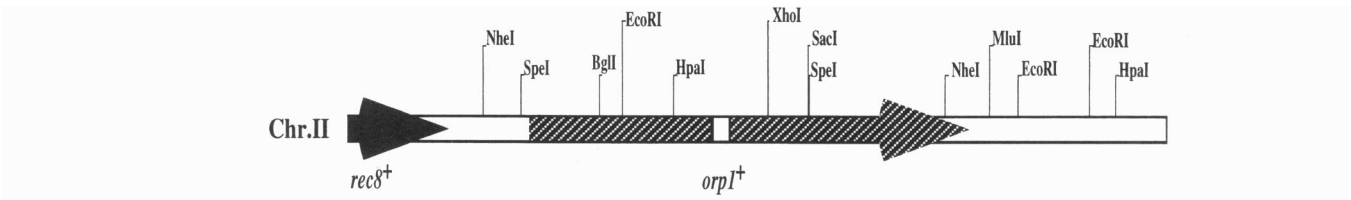
**Gene Disruption.** An *Nhe* I fragment of *orp1*$^+$ (nucleotides −207 to 2108 in Fig. 1), containing almost all of the open reading frame, was replaced by the 1800-bp *ura4*$^+$ gene. A 4383-bp linear *Bam*HI–*Hpa* I fragment containing this *orp1*$^+$::*ura4*$^+$ replacement and the flanking genomic sequences (see map in Fig. 1) was introduced into a diploid strain (*h*$^+$/*h*$^-$ *leu1-32/leu1-32 ura4-D18/ura4-D18 ade6-M210/ade6-M216*). Uracil prototophs were selected and screened to identify *orp1*Δ/*orp1*$^+$ heterozygous diploids. Replacement of the *orp1*$^+$ sequence with *ura4*$^+$ was confirmed by Southern blot analysis of several restriction enzyme digests of DNA derived from transformants. For analysis of the phenotype of haploids with a disruption of the *orp1*$^+$ gene, tetrad analysis was performed by standard methods (17). Large-scale preparation of spores, flow cytometry, and DAPI (4′,6-diamidino-2-phenylindole) staining were performed as described (15).

**Cell Synchronization.** Synchronization was accomplished by releasing a *cdc25-22* strain from a cell cycle block in late $G_2$ (21). For this purpose, an exponentially growing culture of the *cdc25-22* strain was arrested for 4 hr at the nonpermissive temperature of 35.5°C. The arrested culture was then shifted to the permissive temperature of 25°C, and samples were taken at the indicated time points. Preparation of RNA from synchronized cells and Northern blotting were performed as described (15).

## RESULTS

**Identification of *orp1*$^+$.** As part of an effort to identify additional *Sch. pombe* genes involved in regulating DNA replication, we made use of degenerate PCR methods to search for gene products with amino acid sequence homology to Cdc18 protein. The design of appropriate primers was based upon sequence motifs conserved in Cdc18, *S. cerevisiae* Cdc6, and several related sequences that were present in public sequence data bases (see *Materials and Methods*). This approach resulted in the identification of a single PCR product, distinct from *cdc18*$^+$, which was used as a probe to isolate the *orp1*$^+$ gene from an ordered library of *Sch. pombe* genomic clones in cosmids (18, 19). Partial cDNA clones containing the coding sequence of *orp1*$^+$ were also isolated. The *orp1*$^+$ gene is located in cosmid 2G5c (18), which maps just centromere-proximal to the *rad13*$^+$ gene on the long arm of chromosome II.

Sequence analysis of genomic and cDNA clones revealed that *orp1*$^+$ gene contains a single intron and an open reading frame of 707 codons capable of encoding a hypothetical protein of 80 kDa (Fig. 1). The deduced amino acid sequence of p80$^{orp1}$ is 31% identical to the sequence of p65$^{cdc18}$ and contains the A and B components of a classical purine




```
-477 acaatcaaaccacgaactaacaaaacaatagcactagcaactatagtctgcatttttagttaaacaggataaaccatactctgaaattagtgtctctttaaaccaaaaagcaccgatgccatttgaaactttcacatgttaaacttaaaatttttaaa
-318 acgttgaatttgggaattaaacccttgttgaagttgtcaaaatgattgactatgtaataacatttaaatagtatagtataacactttctttagattaatgaatgataatgctagcggataaataattaaaattaattaactacttaccaaatttatct
-159 aattattatatcttaaaattgtggtatattatttgttggtgttgaacatagcatgacagtagcctctatagtatacgagtcttgttagtccagcactacaactcaggataactttgaccattgcaatgttgataaactagtgttgaacttctcttaat
   1 ATG CCT AGA AGA AAG TCA TTG AGG AGT CAA CTA TTA ATT AAC GGC ATT GAT AAA AGT CTG CTA TCT GAT GAC AGC GCT GAC AGT TCT GAT ATT GAC GAA GAG GAA GTT TAC GGT GTT TGG
     M  P  R  R  K  S  L  R  S  Q  L  L  I  N  G  I  D  K  S  L  L  S  D  D  S  A  D  S  S  D  I  D  E  E  E  V  Y  G  V  W
 121 ACT.GAA GAG CCC TTT CAA AAA GAG GCT GGA CGT TCT TAT TAC AGA TCT TTA AAG AAA AAC GAT GTA ATA TAT CGC GTT GGA GAT GAT ATT ACT GTA CAT GAT GGA GAC TCA AGC TTT TAT
     T  E  E  P  F  Q  K  E  A  G  R  S  Y  Y  R  S  L  K  K  N  D  V  I  Y  R  V  G  D  D  I  T  V  H  D  G  D  S  S  F  Y
 241 CTG GGG GTA ATT TGT AAA TTG TAC GAA AAA GCA ATT GAT AAG CAT TCT GGA AAG AAA TAT GTT GAA GCA ATT TGG TAT AGT CGA GCT TAT GCT AAG AGA ATG GAA ATT AAA CCT GAA TAT
     L  G  V  I  C  K  L  Y  E  K  A  I  D  K  H  S  G  K  K  Y  V  E  A  I  W  Y  S  R  A  Y  A  K  R  M  E  I  K  P  E  Y
 361 TTG TTG CCA GAC CGG CAT ATA AAT GAG GTG TAC GTT TCT TGT GGC CGG GAT GAA AAC CTG ACT TCA TGT ATA ATA GAG CAT TGT AAT GTC TAC TCT GAA GCA GAG TTT TTT TCA AAA TTT
     L  L  P  D  R  H  I  N  E  V  Y  V  S  C  G  R  D  E  N  L  T  S  C  I  I  E  H  C  N  V  Y  S  E  A  E  F  F  S  K  F
 481 CCC GCT GGA ATT CCT ACA AAA CGA AAA GAT TTG TTT CCT TGT AAC TTC TTT ATC CGA CGC GGT GTA CAC TTG AAA GTG AAC AAA TAC ACA GAA CCT CTC GAT TGG TCT TAT TAT GCT CAT
     P  A  G  I  P  T  K  R  K  D  L  F  P  C  N  F  F  I  R  R  G  V  H  L  K  V  N  K  Y  T  E  P  L  D  W  S  Y  Y  A  H
 601 AAT CTT GAA AGG ATA GAA GAT CTT TTG GTT GAG ATG GAA GAA AAT TTG CGA CCA ACT AAA AAG AAA TCT GGT TCT AGA GGT CGT GGT CGC CCT CGT AAA TAT CCT TTA CCA AAT GTC GAA
     N  L  E  R  I  E  D  L  L  V  E  M  E  E  N  L  R  P  T  K  K  K  S  G  S  R  G  R  G  R  P  R  K  Y  P  L  P  N  V  E
 721 AGC AAA GAA AGC AGT TCC AAA GTT AAC TCT AAG GAT GAA AAT TTT GAT TTA CAA GAT GAT AGT GAA TCT TCA GAA GAT AAT TTG ACT ATA CAA CCT CAG ACA CCA AGG CGC CGT CAT AAA
     S  K  E  S  S  S  K  V  N  S  K  D  E  N  F  D  L  Q  D  D  S  E  S  S  E  D  N  L  T  I  Q  P  Q  T  P  R  R  R  H  K
 841 AGA TCA AGA CAC AAT TCA TCA AAT TTG GCT TCT ACT CCA AAA AGA AA gtaagttattttatattctaagttttttttcgtatactaacgttgatttag T GGC TAC AAA CAA CCA TTA CAA ATT ACT CCG
     R  S  R  H  N  S  S  N  L  A  S  T  P  K  R  N                                                    G  Y  K  Q  P  L  Q  I  T  P
 971 CTA CCT ATT CGT ATG CTG TCC CTT GAG GAG TTT CAG GGT TCT CCT CAT AGA AAA GCT AGG GCT ATG CTT CAT GTT GCT TCA GTT CCA AGC ACA TTA CAA TGT CGC GAT AAC GAA TTT TCT
     L  P  I  R  M  L  S  L  E  E  F  Q  G  S  P  H  R  K  A  R  A  M  L  H  V  A  S  V  P  S  T  L  Q  C  R  D  N  E  F  S
1091 ACC ATA TTT TCG AAC TTA GAA AGT GCC ATT GAA GAA GAG ACA GGG GCT TGT CTC TAT ATA TCT GGT ACG CCG GGA ACA GGA AAA ACT GCT ACT GTT CAC GAA GTA ATT TGG AAT CTT CAG
     T  I  F  S  N  L  E  S  A  I  E  E  E  T  G  A  C  L  Y  I  S  G  T  P  G  T  G  K  T  A  T  V  H  E  V  I  W  N  L  Q
1211 GAA TTA TCT CGA GAA GGA CAA CTT CCT GAA TTT TCA TTC TGC GAA ATT AAT GGA ATG CGT GTA ACC AGT GCA AAC CAG GCA TAT TCT ATT CTC TGG GAA TCT TTG ACG GGT GAA AGA GTT
     E  L  S  R  E  G  Q  L  P  E  F  S  F  C  E  I  N  G  M  R  V  T  S  A  N  Q  A  Y  S  I  L  W  E  S  L  T  G  E  R  V
1331 ACT CCA ATC CAT GCA ATG GAC CTT CTT GAT AAC CGA TTT ACT CAT GCT TCT CCA AAC CGC AGT AGT TGT GTT GTT CTT ATG GAT GAG CTC GAT CAA CTA GTC ACC CAT AAT CAA AAA GTT
     T  P  I  H  A  M  D  L  L  D  N  R  F  T  H  A  S  P  N  R  S  S  C  V  V  L  M  D  E  L  D  Q  L  V  T  H  N  Q  K  V
1451 TTA TAC AAT TTT TTC AAT TGG CCG TCT CTA CCA CAT TCA CGG TTA ATC GTT GTT GCA GTT GCT AAT ACG ATG GAC TTA CCT GAA CGT ATT TTA TCA AAT CGC ATT TCA TCA CGT TTA GGT
     L  Y  N  F  F  N  W  P  S  L  P  H  S  R  L  I  V  V  A  V  A  N  T  M  D  L  P  E  R  I  L  S  N  R  I  S  S  R  L  G
1571 TTG TCC AGA GTT CCG TTT GAG CCT TAT ACG CAT ACT CAG CTA GAA ATA ATA ATC GCT GCC CGT TTG GAG GCT GTT CGG GAT GAC GAT GTT TTT TCT TCA GAT GCA ATT CGG TTT GCA GCT
     L  S  R  V  P  F  E  P  Y  T  H  T  Q  L  E  I  I  I  A  A  R  L  E  A  V  R  D  D  D  V  F  S  S  D  A  I  R  F  A  A
1691 CGA AAA GTA GCT GCG GTT AGC GGT GAT GCT AGA AGA GCC CTT GAT ATA TGT CGT CGT GCG TCA GAG CTT GCT GAA AAC AAA AAC GGC AAA GTT ACA CCT GGA TTA ATT CAT CAA GCA ATT
     R  K  V  A  A  V  S  G  D  A  R  R  A  L  D  I  C  R  R  A  S  E  L  A  E  N  K  N  G  K  V  T  P  G  L  I  H  Q  A  I
1811 TCC GAA ATG ACA GCT TCA CCG CTT CAA AAA GTA TTA CGA AAT CTC TCA TTC ATG CAG AAA GTA TTT TTA TGT GCT ATA GTC AAT CGT ATG CGC CGG TCT GGA TTT GCA GAG TCG TAT GTT
     S  E  M  T  A  S  P  L  Q  K  V  L  R  N  L  S  F  M  Q  K  V  F  L  C  A  I  V  N  R  M  R  R  S  G  F  A  E  S  Y  V
1931 TAT GAA GTA CTT GAA GAA GCT GAA CGG TTG TTG CGA GTC ATG ACT ACT CCT GAT GCT GAA GCA AAA TTT GGC GAG TTA ATA TTG AGA AGA CCA GAG TTT GGA TAT GTT TTA TCA AGT CTA
     Y  E  V  L  E  E  A  E  R  L  L  R  V  M  T  T  P  D  A  E  A  K  F  G  E  L  I  L  R  R  P  E  F  G  Y  V  L  S  S  L
2051 AGC GAG AAT GGT GTT CTC TAC CTT GAA AAT AAA AGT AGT AGG AAT GCA AGA GTA CGG CTA GCA ATT GCA GAT GAT GAG ATT AAA TTG GCA TTT CGT GGA GAT TCG GAA CTT GCT GGG ATA
     S  E  N  G  V  L  Y  L  E  N  K  S  S  R  N  A  R  V  R  L  A  I  A  D  D  E  I  K  L  A  F  R  G  D  S  E  L  A  G  I
2171 GCA TAA aagctatacttttttggatgaaataggcaatttaccgattgaacaaagtataaaaacttccttaccttacctcttgaatttaaaatgtttacttctaattataaattacgacttaaattatctttttaatttgcccatgataacgaagctg
     A  *
2328 taaacgcgtgttttattcgatactactacttctattattactacctatagagaaccgggtgacgatacttattgtgttatctagtaaagtaaaagagaagtaatagctactgattaaccttagttgtaaaatttcaaaaattcgaattcttacttagcc
2487 ttcctagtttatatagtgcacccatttttcatattcagacttttatgaagcgaatcgaattttgcgttactaacaaaacgcg  2566
```

FIG. 1. Sequence of *orp1*$^+$. The *orp1*$^+$ gene was localized to chromosome II by hybridization to an ordered cosmid library (18, 19). (*Upper*) The hatched arrow indicates the *orp1*$^+$ open reading frame. (*Lower*) Lowercase letters refer to noncoding sequences. A single intron, identified by sequencing the complete cDNA clone, is underlined. Sequence analysis resulted in the identification of *rec8*$^+$ (20) as the gene immediately upstream of *orp1*$^+$. The black arrow in *Upper* indicates the C terminus of the *rec8*$^+$ open reading frame.

nucleotide-binding motif (22). Other features of note in the Orp1 open reading frame are several Cdc2 consensus phosphorylation sites and a highly charged segment with alternating acidic and basic regions that is located N-terminal to the "A" consensus sequence of the purine nucleotide-binding motif.

*Saccharomyces cerevisiae* contains two known proteins with amino acid sequence homology to *Sch. pombe* Cdc18. One of these is Cdc6 and the other is a hypothetical protein deduced from the sequence analysis of chromosome XIII (cosmid 9745; accession no. Z38114). It has been demonstrated that the latter sequence corresponds to *ORC1*, the gene encoding the largest subunit of the ORC of budding yeast (B. Stillman, personal communication). Of the two *S. cerevisiae* proteins, Orp1 is most similar to Orc1 (30% identity), while the Cdc18 protein is most similar to Cdc6 (31% identity). Given the functional similarities of *cdc18*$^+$ and *CDC6*, it seems reasonable to expect that *orp1*$^+$ and *ORC1* mediate analogous functions. Orc1 also has amino acid sequence homology to Sir3, a protein involved in transcriptional silencing (B. Stillman, personal communication). The homology is particularly striking within the N-terminal regions of Orc1 and Sir3. Homology between Orp1 and Sir3 is less evident, although the two proteins are probably distantly related (see below).

**Characterization of *orp1*$^+$.** Using standard methods, we generated a diploid strain in which one copy of the *orp1*$^+$ open reading frame was disrupted with the *ura4*$^+$ gene (17). Analysis of tetrads derived from this strain indicated that *orp1*$^+$ is an essential gene. All 20 tetrads dissected gave rise to 2 viable and 2 inviable spores. The viable spores were all Ura$^-$ as expected for cells carrying the wild-type allele of *orp1*$^+$. The inviable *ura4*$^+$ spores could be rescued by transformation of the diploid with an ARS (autonomously replicating sequence) plasmid carrying either a 6.0-kb *Bam*HI genomic fragment of *orp1*$^+$ or the *orp1*$^+$ cDNA expressed under the control of the *nmt1*$^+$ promoter. To further characterize the phenotype of cells lacking *orp1*$^+$, we studied the fate of the germinating *orp1*$^-$ spores by flow cytometry and microscopy (Fig. 2). The *orp1*$^+$/*orp1*::*ura4*$^+$ diploid was sporulated, and the residual vegetative cells were removed by digestion with Helicase (IBF Biotechnics, Villeneuve-la-Garenne, France; ref. 17). The spores were then inoculated into minimal medium lacking uracil. Under these conditions, only the *orp1*::*ura4*$^+$ spores were able to germinate and grow (17). By 5 hr after inoculation, all of the spores had a 1C DNA content by flow cytometry (Fig. 2 *Upper*). Between 5 and 10 hr after inoculation, the germinating *orp1*$^+$::*ura4*$^+$ spores synthesized DNA and achieved a 2C DNA content. The ungerminated *ura*$^-$*orp1*$^+$ spores remained at a 1C DNA content. The ratio of cells with 2C DNA content to those with 1C DNA content increased at later time points, suggesting that at least some of the spores lacking *orp1*$^+$ are capable of undergoing more than one round of DNA replication and cell division. However, cell division eventually ceased, and the flow cytometry profile became very broad. Microscopic examination revealed that the germinated *orp1*::*ura4*$^+$ spores became increasingly elongated beginning at about 12 hr of incubation (Fig. 2 *Lower*). In addition at later times (15–19 hr), the nuclear morphology became abnormal, often undergoing fragmentation, as illustrated in the inset in Fig. 2 *Lower*. A number of cells with abnormal partitioning of DNA at cell division ("cut" cells) were also observed, suggesting possible abnormalities in the checkpoint mechanism that links mitosis to completion of DNA replication.

As a control, a population of spores derived from a *ura4*$^-$/*ura4*$^+$ *orp1*$^+$/*orp1*$^+$ diploid were inoculated into the same medium. Again, the germinating *ura*$^+$ spores underwent DNA replication between 5 and 10 hr after inoculation although, for reasons not yet understood, the first S phase reproducibly occurred later than in the *orp1*$^+$::ura4$^+$ spores. The cells continued to multiply, so that by 15 hr after inoculation, a majority of the cells had the 2C DNA content characteristic of

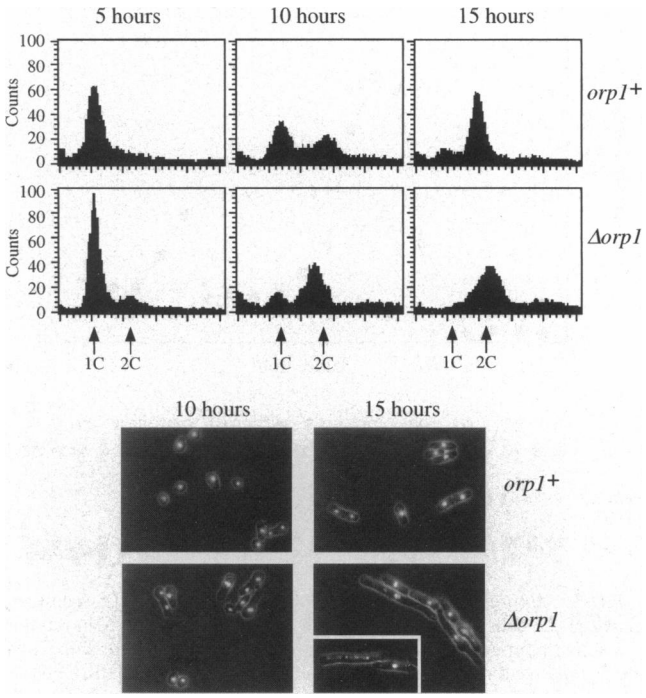



FIG. 2. Phenotype of Δ*orp1* cells. Spores derived from two different diploid strains were germinated in minimal medium lacking uracil. At the indicated times, samples were removed and fixed in ethanol. Fixed cells were rehydrated and stained either with propidium iodide for flow cytometry (*Upper*) or with 4′,6-diamidino-2-phenylindole for fluorescence microscopy (*Lower*). Strains: *orp1*$^+$, *ura4*$^+$/*ura4-D18* *orp1*$^+$/*orp1*$^+$; Δ*orp1*, *orp1*$^+$/*orp1*::*ura4-D18*. The 15-hr Δ*orp1* panel in *Lower* represents the more frequent phenotypes. (*Lower Inset*) Δ*orp1* cells with fragmented nuclei. 1C, 1 genome equivalent (haploid); 2C, 2 genome equivalents (diploid).

vegetatively growing *Sch. pombe* (Fig. 2 *Upper*). They also displayed a normal size distribution and nuclear morphology by microscopy (Fig. 2 *Lower*).

As a final control we examined the fate of spores derived from a *cdc18*$^+$/*cdc18*::*ura4*$^+$ diploid. As previously reported, the *cdc18*::*ura4*$^+$ spores failed to synthesize DNA. The cells remained at a 1C DNA content throughout the experiment, although the 1C peak broadened somewhat over the 15-hr course. Microscopic examination showed that most of the *cdc18*::*ura4*$^+$ spores formed septa and attempted cell division even though DNA replication had not occurred (data not shown). This observation is consistent with the previous suggestion that *cdc18*$^+$ is essential both for initiation of DNA replication and for the replication checkpoint mechanism. Thus, the terminal phenotype of cells lacking *orp1*$^+$ is significantly different from that of cells lacking *cdc18*$^+$.

**Expression of *orp1*$^+$.** The *cdc18*$^+$ gene is expressed during a narrow window of the cell cycle near the $G_1$/S boundary. To examine the expression of *orp1*$^+$ during the cell cycle, we synchronized a population of *cdc25-22* cells by a temperature shift protocol (Fig. 3). At the restrictive temperature, the *cdc25* strain arrested in late $G_2$. Upon shift to the permissive temperature, the cells reentered the cell cycle in relatively synchronous fashion as judged by the septation index. The rise in septation index is an indicator of the transition from the $G_1$ to the S phase of the cell cycle. As a second indicator of the cell synchrony, we examined the level of cdc22$^+$ mRNA following shift to the permissive temperature. The *cdc22*$^+$ gene encodes a subunit of ribonucleotide reductase and is expressed periodically under control of the Cdc10/Sct1 transcriptional activator (23). As expected, the cdc22$^+$ mRNA reached maximal levels at about the time of the $G_1$/S transition. In contrast,

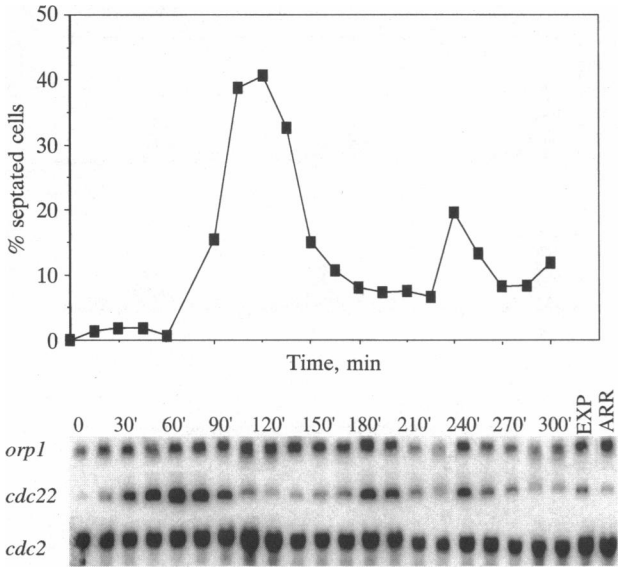



FIG. 3. Expression of orp1$^+$ mRNA during the cell cycle. A culture of *cdc25-22* cells was arrested in $G_2$ by incubation for 5 hr at the restrictive temperature (35.5°). At zero time the cells were shifted to the permissive temperature (25°C) and allowed to proceed through the cell cycle. Samples were removed for isolation of RNA every 15 min. RNA was also prepared from samples obtained immediately before (lane EXP) and after (lane ARR) the incubation at 36°C. (*Upper*) Septation index. (*Lower*) Northern blots of mRNA hybridized to probes specific for *orp1*$^+$, *cdc22*$^+$, and *cdc2*$^+$, of which the last-named served as loading controls and indicated that the samples at 210 min and 225 min were underloaded.

Northern blot analysis indicated that the 2.4-kb orp1$^+$ mRNA is present at roughly the same level throughout the cell cycle.

**The *cdc18*$^+$*/CDC6* Gene Family.** A search of the public sequence data bases has revealed a number of additional sequences that have significant sequence similarity to *cdc18*$^+$ and *CDC6*. Together, these sequences define a novel family of genes. The defining characteristic of the family is a 250-amino acid domain that contains four blocks of conserved sequence elements (Fig. 4). Two of these elements correspond to the A and B motifs characteristic of purine nucleotide-binding domains. The remaining two elements, especially block IV, are unique to the family. Not shown in Fig. 4 are several partial sequences in the data base of expressed sequence tags (dBEST) that encode proteins with amino acid similarity to Cdc18 and Cdc6. These include cDNA fragments obtained from *Plasmodium falciparum* (accession no.: T18071), *Oryza sativa* (rice; accession no.: D41200) and a human placental cDNA library (accession no.: T96858). Thus, the *cdc18*$^+$*/CDC6* family includes members from both the plant and animal kingdoms and from mammals as well as fungi. The hypothetical peptide encoded by the putative human sequence fragment is 49% identical to the corresponding region of *Sch. pombe* Orp1 (Fig. 5). Interestingly, a protein encoded by an archaebacterial plasmid also displays significant homology to the family within the conserved sequence blocks (Fig. 4). Thus, the family appears to be of ancient origin and may be conserved from archaebacteria to man. The Sir3 protein of budding yeast, which is required for transcriptional silencing, appears to be a distant relative of the family and probably shares a common origin (Fig. 4). Although Sir3 exhibits detectable homology to other family members throughout the 250-amino acid conserved domain, it does not contain "A" and "B" purine nucleotide-binding motifs that conform to the consensus.

## DISCUSSION

The *orp1*$^+$ gene was identified by homology to a group of genes that includes *Sch. pombe cdc18*$^+$ and *S. cerevisiae CDC6* and *ORC1*. The protein encoded by *orp1*$^+$ is most closely related to budding yeast Orc1, exhibiting 48% amino acid sequence identity over the segment shown in Fig. 4. Orc1 contains an extended N-terminal region with a very high amino acid sequence similarity to the budding yeast Sir3 protein. Orp1 is smaller than Orc1 and lacks this extended Sir3 homology.

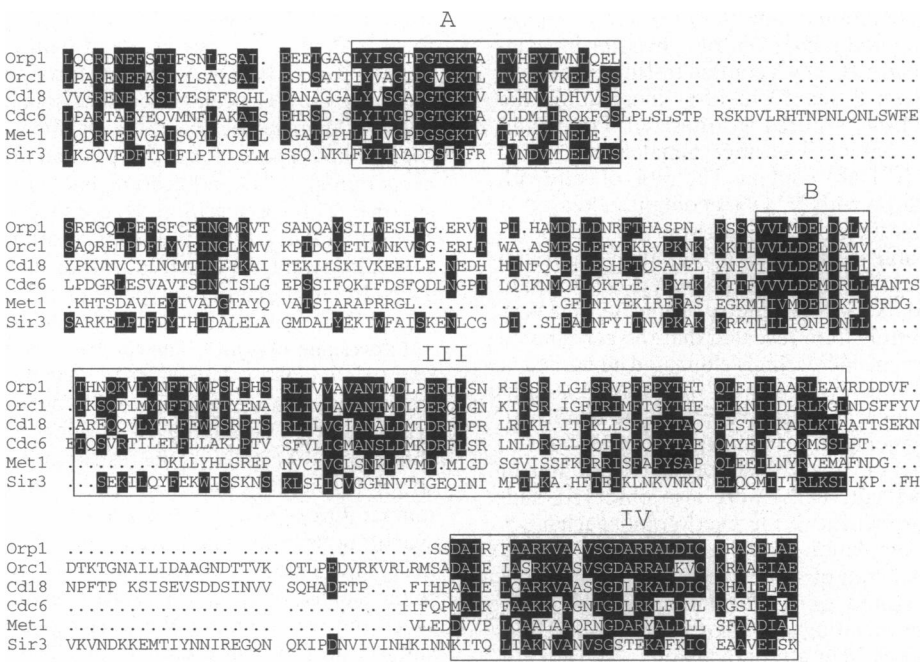



FIG. 4. The *cdc18/CDC6* gene family. Residues highlighted by reverse contrast were present in at least three of the six family members shown. Orp1, *Sch. pombe* Orp1 protein (residues 239–472); Orc1, *S. cerevisiae* Orc1 protein (residues 449–718); Cd18, *Sch. pombe* Cdc18 protein (accession no. A40726; residues 169–434); Cdc6, *S. cerevisiae* Cdc6 protein (accession no. X65299; residues 78–346); Met, *Methanobacterium thermoformicicum* plasmid pFV1 hypothetical protein (accession no. P29569; residues 24–240); Sir3, *S. cerevisiae* Sir3 protein (accession no. P06701; residues 549–814).

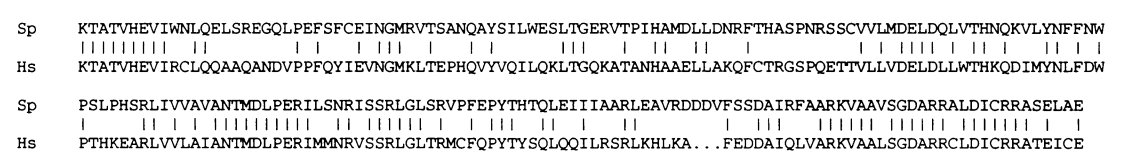



FIG. 5. Alignment of Orp1 and a putative human homolog. The cDNA clone containing the putative human homologue of Cdc18/Cdc6 was provided by the I.M.A.G.E. (Integrated Molecular Analysis of Genomes and their Expression) Consortium. The clone contained an open reading frame of 82 residues and was extended by PCR with degenerate primers as described in text. Solid lines indicate identical amino acids. Hs, *Homo sapiens*; Sp, *Sch. pombe.*

Nevertheless, the similarity of the two proteins is striking and strongly suggests that *orp1*$^+$ encodes a component of an *Sch. pombe* origin recognition complex. Origins of DNA replication have been isolated and characterized in *Sch. pombe*, but the proteins that bind to them have not been identified. The identification of *orp1*$^+$ may facilitate the detection of such proteins.

The *orp1*$^+$ gene is essential for viability, and germinating spores that lack *orp1*$^+$ have an interesting phenotype. They are capable of initiating DNA replication at least once, but eventually they elongate with abnormal nuclei and cease to divide. If *orp1*$^+$ encodes a component of an essential ORC in *Sch. pombe*, our data suggest that the mutant spores contain sufficient functional Orp1 protein to initiate DNA replication leading to duplication of at least a large fraction of the chromosomal DNA. This suggestion is in keeping with the observation that the *S. cerevisiae* ORC complex is probably present at origins of replication throughout the cell cycle (12). Thus, the chromosomal origins of replication in spores derived from *orp1*$^+$/*orp1::ura4* diploids may already contain a previously bound ORC complex containing functional Orp1 protein. This situation might allow a few rounds of DNA synthesis, but the mutant cells would eventually die because no new Orp1 protein can be synthesized.

It is interesting that *Sch. pombe* contains two related genes probably involved in DNA replication, one of which, *orp1*$^+$, is expressed continuously, and the other, *cdc18*$^+$, is expressed periodically during the cell cycle. It has been reported that the budding yeast *CDC6* gene is expressed periodically (24, 25). The expression pattern of *ORC1* has not yet been reported, but *in vivo* footprinting data indicate that the ORC complex is present throughout the cell cycle (12). Recent genetic and biochemical studies indicate that the Cdc6 protein interacts with ORC (26). Thus, the periodically expressed *cdc18*$^+$/*CDC6* genes may play key roles in the machinery that determines the timing of initiation of DNA replication. Given the similarities among these proteins, it is possible that the function of Cdc18/Cdc6 is mediated via interactions with the ORC complex at origins of DNA replication. For example, Cdc18 and Cdc6 proteins might interact with subunits of ORC via interactions similar to those utilized by Orp1 and Orc1, possibly even replacing the latter subunits prior to S phase. The availability of the genes should make it possible to test speculations such as these directly.

Whatever their biochemical roles, it is likely that both *cdc18*$^+$ and *orp1*$^+$ are involved in critical steps required for DNA replication. The identification of a family of related genes conserved from archaebacteria to man also attests to their fundamental importance. If, as seems likely, all of the members of the *cdc18/CDC6* family encode regulators of DNA replication, the identification and characterization of the human member may provide a way to approach the problem of understanding initiation of DNA replication in higher eukaryotes.

We thank the members of the Kelly lab for stimulating discussions; P. Simanceck for valuable help; and Dr. D. Herendeen, D. Bellows, and Dr. B. Stillman for communicating results prior to publication. This work was supported by grants from the National Institutes of Health.

1. Kelly, T. J., Nurse, P. & Forsburg, S. L. (1993) *Cold Spring Harbor Symp. Quant. Biol.* **58,** 637–644.
2. Campbell, J. L. & Newlon, C. S. (1991) in *Chromosomal DNA Replication*, eds. Broach, J. R., Pringle, J. R. & Jones, E. W. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 1, pp. 41–146.
3. Coverley, D. & Laskey, R. A. (1994) *Annu. Rev. Biochem.* **63,** 745–776.
4. DePamphilis, M. L. (1993) *J. Biol. Chem.* **268,** 1–4.
5. Stillman, B. (1994) *J. Biol. Chem.* **269,** 7047–7050.
6. Johnston, L. H. & Lowndes, N. F. (1992) *Nucleic Acids Res.* **20,** 2403–2410.
7. Diffley, J. F. (1994) *Curr. Opin. Cell Biol.* **6,** 368–372.
8. Bell, S. P., Kobayashi, R. & Stillman, B. (1993) *Science* **262,** 1844–1849.
9. Foss, M., McNally, F. J., Laurenson, P. & Rine, J. (1993) *Science* **262,** 1838–1844.
10. Li, J. J. & Herskowitz, I. (1993) *Science* **262,** 1870–1874.
11. Micklem, G., Rowley, A., Harwood, J., Nasmyth, K. & Diffley, J. F. (1993) *Nature (London)* **366,** 87–89.
12. Diffley, J. F., Cocker, J. H., Dowell, S. J. & Rowley, A. (1994) *Cell* **78,** 303–316.
13. Caligiuri, M. & Beach, D. (1993) *Cell* **72,** 607–619.
14. Reymond, A., Marks, J. & Simanis, V. (1993) *EMBO J.* **12,** 4325–4334.
15. Kelly, T. J., Martin, G. S., Forsburg, S. L., Stephen, R. J., Russo, A. & Nurse, P. (1993) *Cell* **74,** 371–382.
16. Hogan, E. & Koshland, D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 3098–3102.
17. Moreno, S., Klar, A. & Nurse, P. (1991) *Methods Enzymol.* **194,** 795–823.
18. Hoheisel, J. D., Maier, E., Mott, R., McCarthy, L., Grigoriev, A. V., Schalkwyk, L. C., Nizetic, D., Francis, F. & Lehrach, H. (1993) *Cell* **73,** 109–120.
19. Maier, E., Hoheisel, J. D., McCarthy, L., Mott, R., Grigoriev, A. V., Monaco, A. P., Larin, Z. & Lehrach, H. (1992) *Nat. Genetics* **1,** 273–277.
20. Lin, Y., Larson, K. L., Dorer, R. & Smith, G. R. (1992) *Genetics* **132,** 75–85.
21. Moreno, S., Hayles, J. & Nurse, P. (1989) *Cell* **58,** 361–372.
22. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) *EMBO J.* **1,** 945–951.
23. Lowndes, N. F., McInerny, C. J., Johnson, A. L., Fantes, P. A. & Johnston, L. H. (1992) *Nature (London)* **355,** 449–453.
24. Zwerschke, W., Rottjakob, H. W. & Kuntzel, H. (1994) *J. Biol. Chem.* **269,** 23351–23356.
25. Zhou, C. & Jong, A. (1990) *J. Biol. Chem.* **265,** 19904–19909.
26. Liang, C., Weinreich, M. & Stillman, B. (1995) *Cell* **81,** 667–676.