permeabilized in 0.5% Nonidet P-40 in phosphate-buffered saline, and labeled with mouse mAb 12CA5 followed by FITC-labeled goat antibody to mouse IgG (Jackson ImmunoResearch Labs). DNA was stained with DAPI (0.2 μg/ml). Images of interphase nuclei were obtained with a Zeiss Axioplan microscope equipped with a Kodak DCS-200 digital camera. Images were noise-filtered with a 3 × 3 median filter, corrected for background with nontransfected cells as a reference, and merged to obtain triple labeling images using Adobe Photoshop. Chromosome spreads were photographed on Kodak Gold II 400 ISO film and converted to digital images with a Nikon slide scanner, after which FITC and DAPI images were corrected for background and superimposed.
16. T. de Lange *et al., Mol. Cell. Biol.* **10**, 518 (1990).
17. B. van Steensel and T. de Lange, data not shown.
18. T. de Lange, *EMBO J.* **11**, 717 (1992); H. Tommerup, A. Dousmanis, T. de Lange, *Mol. Cell. Biol.* **14**, 5777 (1994); S. Lejnine, V. L. Makarov, J. P. Langmore, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2393 (1995).
19. M. N. Conrad, J. H. Wright, A. J. Wolf, V. A. Zakian, *Cell* **63**, 739 (1990); A. J. Lustig, S. Kurtz, D. Shore, *Science* **250**, 549 (1990); L. Sussel and D. Shore, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7749 (1991); F. Klein *et al., J. Cell Biol.* **117**, 935 (1992); G. Kyrion, K. A. Boakye, A. J. Lustig, *Mol. Cell. Biol.* **12**, 5159 (1992); G. Kyrion, K. Liu, C. Liu, A. J. Lustig, *Genes Dev.* **7**, 1146 (1993); D. Shore, *Trends Genet.* **10**, 408 (1994); V. A. Zakian, in *Telomeres*, E. H. Blackburn and C. W. Greider, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1995), pp. 107–138.
20. M. J. McEachern and E. H. Blackburn, *Nature* **376**, 403 (1995).
21. D. Rhodes, personal communication.
22. D. Kipling and H. J. Cooke, *Nature* **347**, 400 (1990); J. A. Starling, J. Maule, N. D. Hastie, R. C. Allshire, *Nucleic Acids Res.* **18**, 6881 (1990).
23. C. M. Counter *et al., EMBO J.* **11**, 1921 (1992); A. J. Klingelhutz *et al., Mol. Cell. Biol.* **14**, 961 (1994); C. M. Counter *et al., J. Virol.* **68**, 3410 (1994).
24. H. J. Cooke and B. A. Smith, *Cold Spring Harbor Symp. Quant. Biol.* **51**, 213 (1986); R. C. Allshire *et al., Nature* **332**, 656 (1988); C. B. Harley, A. B. Futcher, C. W. Greider, *ibid.* **345**, 458 (1990); N. D. Hastie *et al., ibid.* **346**, 866 (1990).
25. N. D. Hastie and R. C. Allshire, *Trends Genet.* **5**, 326 (1989).
26. C. M. Counter, H. W. Hirte, S. Bachetti, C. Harley, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2900 (1994); N. W. Kim *et al., Science* **266**, 2011 (1994); J. Feng *et al., ibid.* **269**, 1236 (1995).
27. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
28. We thank B. Stillman and S. Bell for information about ORC1 and N. Andrews, R. Benezra, A. Bianchi, R. Bose, R. Darnell, J. Feng, R. Lang, M. Lui, A. Lustig, N. Segil, O. Sibon, S. Smith, and J. Young for their contributions and advice. Supported by grants from the Lucille P. Markey Charitable Trust and NIH (GM49046) to T.d.L., grants from the National Cancer Institute (P30 CA08748-29) and NSF (BIR-9420123) to P.T., and a Irma T. Hirschl Career Scientist Award to P.T. B.v.S. is the recipient of a Human Frontier Science Program postdoctoral fellowship.

3 October 1995; accepted 1 November 1995

Conserved Initiator Proteins in Eukaryotes

Kimberley A. Gavin, Masumi Hidaka, Bruce Stillman*

The origin recognition complex (ORC), a multisubunit protein identified in *Saccharomyces cerevisiae*, binds to chromosomal replicators and is required for the initiation of cellular DNA replication. Complementary DNAs (cDNAs) encoding proteins related to the two largest subunits of ORC were cloned from various eukaryotes. The cDNAs encoding proteins related to *S. cerevisiae* Orc1p were cloned from the budding yeast *Kluyveromyces lactis*, the fission yeast *Schizosaccharomyces pombe,* and human cells. These proteins show similarity to regulators of the S and M phases of the cell cycle. Genetic analysis of *orc1*$^+$ from *S. pombe* reveals that it is essential for cell viability. The cDNAs encoding proteins related to *S. cerevisiae* Orc2p were cloned from *Arabidopsis thaliana, Caenorhabditis elegans*, and human cells. The human ORC-related proteins interact in vivo to form a complex. These studies suggest that ORC subunits are conserved and that the role of ORC is a general feature of eukaryotic DNA replication.

The replicon model for the initiation of DNA replication postulates that for DNA synthesis to occur, an initiator protein is required for recognition of a specific replicator sequence in the chromosome (*1*). In this model, recognition of the replicator by the initiator protein determines the location of an origin of DNA replication. Data from studies of prokaryotes and eukaryotic viruses support this hypothesis and suggest that the mechanism of initiation of DNA replication is conserved in eukaryotes. The initiation of DNA replication in eukaryotic cells is tightly controlled during the cell cycle and throughout development to ensure that duplication of the genome occurs only once per cell cycle. Thus, initiation is a key regulatory step in DNA replication.

In eukaryotes, the nature of initiator proteins and replicator elements remains unclear (*2*). Attempts to define precisely the origins of DNA replication have been largely unsuccessful. A notable exception is the yeast *Saccharomyces cerevisiae*, in which the origins of DNA replication have been physically mapped (*3*). Autonomously replicating sequences (ARS) have been characterized at the molecular level, revealing a modular structure (*4*, *5*). All ARS elements contain an essential ARS consensus sequence (ACS) and other elements that together are required for ARS function. In addition to characterized replication origins, an initiator protein complex has been isolated (*6*). ORC was identified by its ability to recognize and bind to the ACS in an adenosine triphosphate (ATP)–dependent manner. ORC recognizes a bipartite sequence within the replicator (*6*, *7*) and is bound to the DNA throughout the cell cycle (*8*). In cooperation with other cell cycle proteins, ORC determines the frequency of initiation in the *S. cerevisiae* genome (*9*–*12*).

ORC is a multisubunit protein complex consisting of six polypeptides with apparent molecular mass ranging from 50 to 120 kD. The genes encoding all the subunits have been cloned (*11*–*14*) and all are essential for cell viability. All six genes encode novel proteins, although *ORC1* and *ORC5* encode proteins that have regions of sequence similarity that are present in previously identified proteins. *ORC1* encodes the largest subunit of ORC, and its protein product is related to two other known yeast proteins (*13*). The protein most related to Orc1p is Sir3p, a protein involved in transcriptional silencing (*15*). Although these proteins are related over their entire length, the most striking similarity is found in the first 220 amino acids. A pairwise comparison shows that the proteins are 50% identical and 63% similar over this region. Orc1p is also related to two cell division cycle (CDC) regulators of initiation of DNA replication and control of the G_2- to M-phase transition, *S. cerevisiae* Cdc6p and *Schizosaccharomyces pombe* Cdc18p (*16*). A 270–amino acid region in Orc1p (residues 449 to 717) contains a purine nucleotide-binding motif (*17*) that is present in both Cdc6p and Cdc18p. In addition to the canonical P-loop and A-loop necessary for nucleotide metabolism (*17*), there are sequences flanking these nucleotide-binding motifs that are present only in Orc1p, Cdc6p, and Cdc18p (*13*). Orc1p is 50% identical to Cdc6p and Cdc18p across this region. We refer to this domain as the CDC–nucleoside triphosphate–binding (CDC-NTP) domain. Although Sir3p shares sequence homology with Orc1p across this domain, key residues essential for nucleotide binding are absent from Sir3p (*13*).

The nature of replicators in metazoan species is not clear, and in *Xenopus* early embryos their very existence has been questioned (*18*). To begin to address whether the replicon model applies to cell chromosome replication in higher eukaryotes, we identified genes related to *ORC1* in humans and other organisms (Fig. 1). We used a polymerase chain reaction (PCR) strategy, with primers based

K. A. Gavin, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA, and Program in Genetics, State University of New York, Stony Brook, NY 11794, USA. M. Hidaka and B. Stillman, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA.

*To whom correspondence should be addressed.

on regions conserved between Orc1p and Sir3p, and identified a gene in the related budding yeast *Kluyveromyces lactis* whose product KlOrc1p is related to Orc1p (*19*). KlOrc1p is 50% identical and 68% similar to *S. cerevisiae* Orc1p (ScOrc1p) overall (Fig. 1B). Within KlOrc1p are two regions that are highly related to ScOrc1p. The NH_2-terminal domain (NTD), with high identity between Sir3p and Orc1p, is conserved in KlOrc1p; the proteins are 32% identical and 43% similar over residues 1 to 220. However, the conservation between KlOrc1p and ScOrc1p is most striking in the CDC-NTP domain (69% identical and 81% similar; Fig. 1A). To identify other *ORC1*-related genes, we used a DNA probe from *S. cerevisiae ORC1* containing the CDC-NTP domain and isolated a related gene from *S. pombe* by low-stringency hybridization (*20*). Sequence analysis revealed that *S. pombe* Orc1p (SpOrc1p) and ScOrc1p are 30% identical overall (Fig. 1B). Like KlOrc1p, SpOrc1p is most related to ScOrc1p across the CDC-NTP domain; this region is 46% identical and 61% similar to ScOrc1p (Fig. 1A). In contrast with KlOrc1p, the NTD of Orc1p and of Sir3p is absent in SpOrc1p. Indeed, the first 200 amino acids of *S. cerevisiae* Orc1p are not essential for viability but are required for mating type gene repression (*13*). The absence of this domain in a species that has evolutionarily diverged from the budding yeasts suggests that specialized mating type control and silencing are not conserved.

The identification of *ORC1*-related genes in both budding and fission yeasts suggests that Orc1p is conserved in higher eukaryotes. To investigate this, we again used PCR to identify a human *ORC1*-related gene. The PCR primers were designed based on conserved residues in the CDC-NTP domain that is present in ScOrc1p, KlOrc1p, and SpOrc1p; and reverse transcription PCR (RT-PCR) reactions were then carried out on human RNA. An amplified DNA fragment similar to *ORC1* was generated and used as a probe to isolate full-length cDNAs (*21*). Sequence analysis revealed that this gene product was related to Orc1p, Cdc6p, and Cdc18p (Fig. 1A). A BLAST database search comparing the human gene product with both Orc1p and Cdc6p or Cdc18p showed the protein to be more closely related to Orc1p ($P = 7.4 \times 10^{-52}$ for Orc1p, compared with $P = 3.6 \times 10^{-10}$ for Cdc6p) (*21*). *Homo sapiens* Orc1p (HsOrc1p) is 27% identical to ScOrc1p overall, but is 44% identical and 60% similar to ScOrc1p across the CDC-NTP domain (Fig. 1A). The Orc1 proteins can be distinguished from Cdc6p and Cdc18p by sequences in the CDC-NTP domain that are conserved among Orc1 proteins but not among Cdc18p or Cdc6p (for example, residues upstream and downstream of CDC-NTP box 4). In addition, there are shorter stretches of conserved residues both upstream of the CDC-NTP domain and downstream, extending to the end of the proteins (Fig. 1B).

Genetic analysis of *ScORC1* shows that the gene is essential for cell viability (*13*). To test whether *S. pombe orc1*$^+$ is essential, we created a null allele of *orc1*$^+$ by replacing a 1.4-kb fragment of the coding sequence with the *ura4*$^+$ gene (*22*). Southern (DNA) blot analysis confirmed that stable Ura$^+$ transformants carried one wild-type allele and one disrupted allele. Sporulation of the heterozygous diploid produced two viable *ura*$^-$ spores and two nonviable spores, demonstrating that *orc1*$^+$ is essential (*23*). The lethal phenotype of the disruption could be rescued by a plasmid containing a genomic fragment of *orc1*$^+$ or an influenza

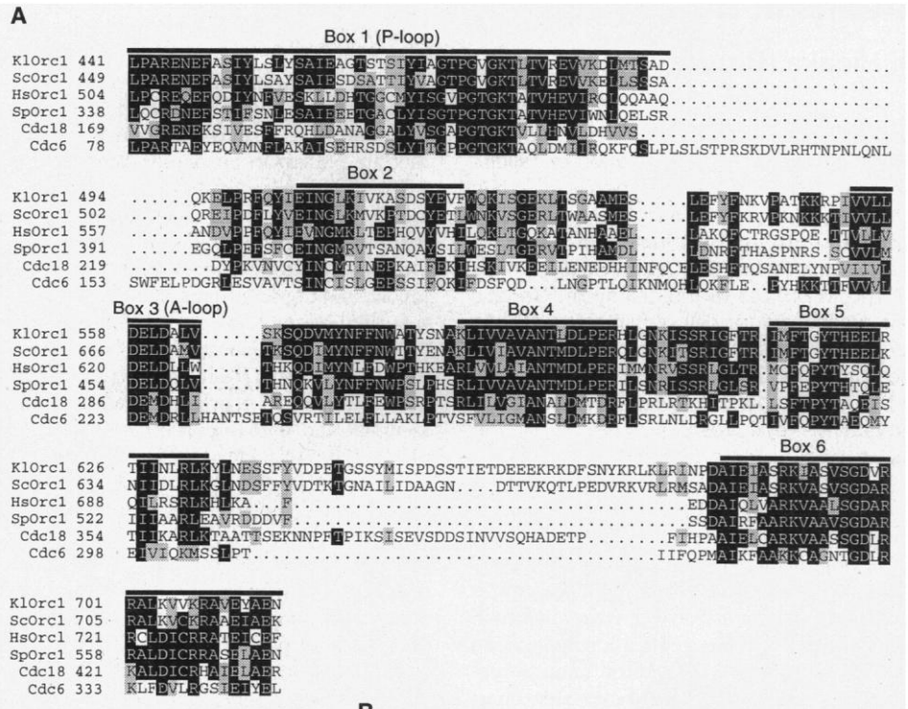

Fig. 1. Orc1-related proteins in diverse organisms (*33*). (**A**) Multiple alignment of Orc1 proteins with Cdc6p and Cdc18p across the CDC-NTP region. KlOrc1, ScOrc1, HsOrc1, and SpOrc1 indicate Orc1 proteins from *K. lactis, S. cerevisiae*, human cells, and *S. pombe*, respectively, and Cdc18 and Cdc6 are Cdc18p from *S. pombe*, and Cdc6p from *S. cerevisiae*, respectively. Identical amino acids conserved among three or more sequences are in black; identical amino acids conserved among two sequences, or similar amino acids conserved among three or more sequences, are shaded gray. Amino acid numbers are shown on the left. Boxes 1 through 6 show regions of high similarity in Orc1p, Cdc18, and Cdc6. (**B**) Comparison of Sir3p; Orc1 proteins from *S. cerevisiae, K. lactis, S. pombe*, and human cells; Cdc6p; and Cdc18p. The NTD conserved among Sir3p, ScOrc1p, and KlOrc1p is indicated by a hatched box. The CDC-NTP motif is indicated by a shaded box, with the integral nucleotide binding site represented by a dark, shaded box. The percent identity of Orc1-related proteins with respect to ScOrc1p is shown on the right. Regions of conservation among Orc1 proteins and Sir3p or Cdc6 and Cdc18p at the COOH-terminus are indicated by black lines; conservation among Orc1 proteins is indicated by gray lines. aa, amino acid.

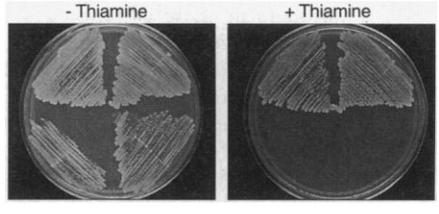

Fig. 2. *Schizosaccharomyces pombe orc1*$^+$ is essential for viability. Wild-type cells (top right quadrant), *orc1::ura4*$^+$ cells containing either a genomic *orc1*$^+$ under its native promoter (top left quadrant), or a plasmid containing the *orc1*$^+$ cDNA under the control of the *nmt1* promoter (bottom) were grown in the absence (−) or presence (+) of thiamine at 30°C. Two independent colonies containing *nmt1-orc1*$^+$ were tested.

virus hemagglutinin epitope–polyhistidine–tagged *orc1*$^+$ cDNA under the control of the inducible *nmt1* promoter (*24*). The *orc1* gene disrupted mutant cells carrying *orc1*$^+$ on a plasmid were grown in the absence or presence of thiamine (*25*). Under nonrepressing conditions (in the absence of thiamine), *nmt1-orc1* cells were viable, but in the presence of thiamine, no viable cells were recovered (Fig. 2). The presence or absence of thiamine had no effect on cells that contained a genomic *orc1*$^+$ under the control of its own promoter. Although *KlORC1* and *SpORC1* are related to *ScORC1*, neither can functionally substitute for *ScORC1* (*23*). Analysis of *S. pombe orc1*$^+$ is in preparation (*26*).

As Orc1p in *S. cerevisiae* is part of a multiprotein complex, the identification of related proteins in diverse eukaryotes suggests that additional ORC proteins are conserved. Database searches revealed expressed sequence tags of partial cDNAs from *A. thaliana* and *C. elegans* that encoded amino acids related to *S. cerevisiae* Orc2p. The partial gene fragments were used as probes to isolate full-length cDNAs from both *A. thaliana* (*27*) and *C. elegans*. (*28*). Comparison of the full-length proteins with *S. cerevisiae* Orc2p revealed short stretches of conservation (Fig. 3). To identify an *ORC2*-related gene in human cells, primers based on these conserved amino acids were used in RT-PCR reactions containing human RNA. DNA fragments that were related to *ORC2* were obtained and used as probes to isolate a full-length cDNA (*29*). Alignment of AtOrc2p, HsOrc2p, CeOrc2p, and ScOrc2p shows that the proteins are 23 to 31% identical overall (Fig. 3). We and others have also isolated a full-length cDNA encoding *Drosophila* Orc2p and shown it to be similar to those we describe here (*23*, *30*). The sizes of the Orc2 proteins are highly variable, and this variability is seen predominantly in the NH_2-termini. The region conserved among the four species is localized to the COOH-terminus. The Orc2 proteins lack any sequence characteristics that are shared with other previously identified proteins. Unlike Orc1 proteins, in which the similarity across the CDC-NTP motif spans 270 amino acids, the regions conserved among the Orc2-related proteins are short and are separated by less well-conserved sequences.

The *S. cerevisiae* Orc1p and Orc2p interact with four additional subunits in vivo to form a functional initiator complex, suggesting that HsOrc1p and HsOrc2p would similarly associate with each other as part of a multiprotein complex. To test this, a T7 epitope–tagged HsOrc1p was transiently overexpressed in human 293 cells (*31*). Rabbit polyclonal antisera raised to glutathione-S-transferase (GST)–fused HsOrc2p were used to immunoprecipitate endogenous HsOrc2p from extracts prepared from transfected cells (*31*). Immunoblot analysis with a monoclonal antibody specific for the epitope detected HsOrc1p in samples containing HsOrc2p antisera (Fig. 4, lanes 5 and 6) but not from immunoprecipitations using preimmune sera (Fig. 4, lane 4) or from extracts of mock-transfected cells (Fig. 4, lane 3). The stronger signals between 30 and 50 kD represent cross-reactivity to immunoglobulins. These data demonstrate that HsOrc1p and HsOrc2p are part of a complex in vivo. Moreover, the endogenous HsOrc1p and HsOrc2p cofractionate during column chromatography (*23*).

Our data demonstrate that two subunits of the yeast initiator complex are conserved in different eukaryotes, including complex vertebrates. Thus, it is likely that other subunits of the protein complex are conserved. Recent identification of a multiprotein complex from *Drosophila* that contains two ORC-related proteins (*30*) further supports this hypothesis. The high degree of conservation of the CDC-NTP domain among the Orc1-related proteins also suggests that the nucleotide requirement for DNA binding is also conserved. Data from *S. cerevisiae ORC1* show that an intact nucleotide-binding mo-

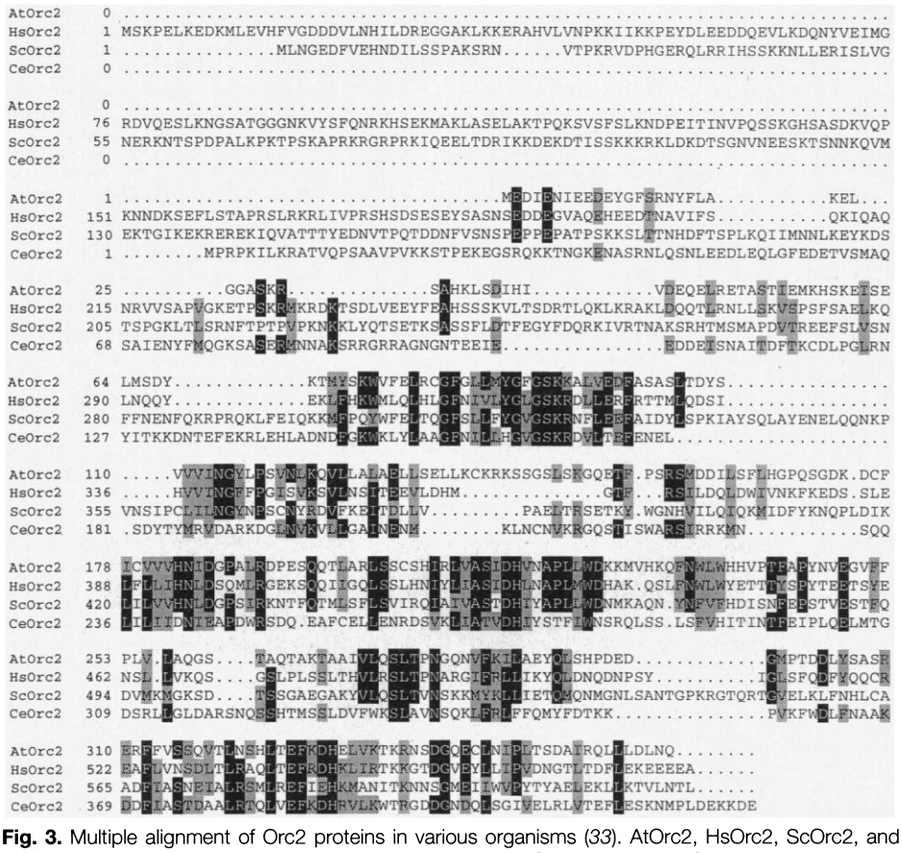

Fig. 3. Multiple alignment of Orc2 proteins in various organisms (*33*). AtOrc2, HsOrc2, ScOrc2, and CeOrc2 represent Orc2 polypeptides of *A. thaliana*, humans, *S. cerevisiae*, and *C. elegans*, respectively. Identical amino acids conserved among three or more species are in black; gray shading indicates similar amino acids conserved among three or more sequences. Amino acid numbers are shown on the left.

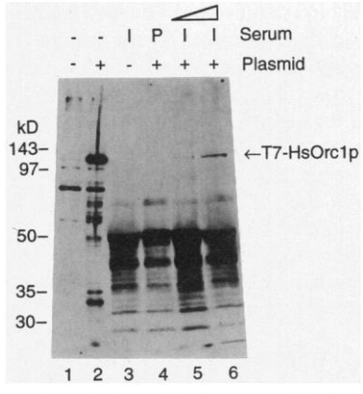

Fig. 4. Association of HsOrc1p and HsOrc2p in vivo. Human 293 cells were transiently transfected with a plasmid expressing T7-HsORC1 under the cytomegalovirus promoter or were mock-transfected. Whole-cell extracts of mock-transfected (lane 1) or transfected (lane 2) cells were prepared and used for immunoprecipitations. Immune complexes were collected, separated by SDS-PAGE, and transferred to nitrocellulose membrane. Immunoblot analysis was done with a monoclonal antibody to the T7 epitope. Lane 3, mock-transfected extract incubated with 4 μl of immune (I) sera. T7-HsORC1 extracts were incubated with 4 μl of preimmune (P) sera (lane 4) or with increasing amounts of antibodies to HsOrc2p (1 μl, lane 5; 4 μl, lane 6), as indicated by the triangle. T7-HsOrc1p is indicated by an arrow.

tif is essential for cell viability (*13*). Moreover, initiator proteins in bacteria, bacteriophages, and viruses commonly show a dependence on nucleo- tide binding for function (*32*). Although the exact role of ATP or GTP in Orc1p function has not yet been determined, the finding that the nucleotide-binding motif is conserved through evolution suggests that its role is important.

Genetic data support the idea that ORC in association with other replication proteins determines the frequency of origin firing in the genome (*10*). These data suggest that the ORC complex is conserved [our data and (*30*)] and that initiation of replication in metazoan species is likely to be DNA sequence–specific. This sequence specificity for establishment of replication origins may be subject to developmental regulation. For instance, the replication of foreign DNAs in *Xenopus* early embryos or extracts is subject to cell cycle regulation; however, initiation occurs in an apparently sequence-independent manner (*18*). If ORC is required for replication in the early embryo, how might this be reconciled with an apparent lack of DNA sequence specificity? Under stringent binding conditions, ORC binds to and requires a bipartite sequence element in yeast replicators (*7*). In contrast, the stringency of ORC binding to the origin DNA is relaxed as the competitor DNA concentration is decreased. By analogy, a high concentration of ORC relative to origin DNA present in an early embryonic cell (*30*) could cause sequence-specific initiation to be relaxed, because ORC would bind to only one of the two recognition sites in the replicator. As the genome replicates and the stockpile of ORC is reduced, this would cause an increased specificity of ORC for high-affinity, bipartite replicator sequences, and consequently cause a decrease in the frequency of initiation of DNA replication in the genome. The isolation of ORC from metazoan species should facilitate the identification of these replicator sequences and further understanding of the regulation of replication initiation.

REFERENCES AND NOTES

1. F. Jacob, S. Brenner, F. Cuzin, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 329 (1963).
2. J. L. Hamlin and P. A. Dijkwel, *Curr. Opin. Genet. Dev.* **5**, 153 (1995); M. L. DePamphilis, *J. Biol. Chem.* **268**, 1 (1993).
3. J. L. Campbell and C. S. Newlon, in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, J. R. Broach, J. R. Pringle, E. W. Jones, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1991), pp. 41–146; A. Rowley, S. J. Dowell, J. F. X. Diffley, *Biochim. Biophys. Acta* **1217**, 239 (1994).
4. S. S. Walker, A. K. Malik, S. Eisenberg, *Nucleic Acids Res.* **19**, 6255 (1991); Y. Marahrens and B. Stillman, *Science* **255**, 817 (1992); J. F. Theis and C. S. Newlon, *Mol. Cell. Biol.* **14**, 7652 (1994).
5. H. Rao, Y. Marahrens, B. Stillman, *Mol. Cell. Biol.* **14**, 7643 (1994).
6. S. P. Bell and B. Stillman, *Nature* **357**, 128 (1992).
7. H. Rao and B. Stillman, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2224 (1995); A. Rowley, J. H. Cocker, J. Harwood, J. F. X. Diffley, *EMBO J.* **14**, 2631 (1995).
8. J. F. X. Diffley and J. H. Cocker, *Nature* **357**, 169 (1992); _____, S. J. Dowell, A. Rowley, *Cell* **78**, 303 (1994).
9. C. A. Fox, S. Loo, A. Dillin, J. Rine, *Genes Dev.* **9**, 911 (1995).
10. C. Liang, M. Weinreich, B. Stillman, *Cell* **81**, 667 (1995).
11. S. P. Bell, R. Kobayashi, B. Stillman, *Science* **262**, 1844 (1993).
12. S. Loo *et al.*, *Mol. Biol. Cell* **6**, 741 (1995).
13. S. P. Bell, J. Mitchell, J. Leber, R. Kobayashi, B. Stillman, *Cell* **83**, 563 (1995).
14. M. Foss, F. J. McNally, P. Laurenson, J. Rine, *Science* **262**, 1838 (1993); J. J. Li and I. Herskowitz, *ibid.*, p. 1870; G. Micklem, A. Rowley, J. Harwood, K. Nasmyth, J. F. X. Diffley, *Nature* **366**, 87 (1993).
15. D. Shore, M. Squire, K. A. Nasmyth, *EMBO J.* **3**, 2817 (1984).
16. A. Bueno and P. Russell, *ibid.* **11**, 2167 (1992); T. J. Kelly *et al.*, *Cell* **74**, 371(1993); S. Piatti, C. Lengauer, K. Nasmyth, *EMBO J.* **14**, 3788 (1995).
17. E. V. Koonin, *Nucleic Acids Res.* **21**, 2541 (1993).
18. R. M. Harland and R. A. Laskey, *Cell* **21**, 761 (1980); M. Méchali and S. Kearsey, *ibid.* **38**, 55 (1984); J. J. Blow and A. M. Sheman, *J. Cell Sci.* **95**, 383 (1990).
19. PCR reactions were carried out with *K. lactis* genomic DNA and degenerate oligonucleotides ORC1–3N (5′-AC[A/C/T]TATTC[A/C/T]GC[A/C/T]TATATGAT[A/C/T] CA-3′) and ORC1–2C(5′-AAGTTC[T/G/A]GC[T/G/A] GT[T/C]AA[G/A]TA[T/C]AA[T/G]TC-3′) based on regions of conservation between Orc1p and Sir3p. A 230–base pair (bp) amplified product related to *S. cerevisiae* Orc1p was used as a probe to screen a *K. lactis* genomic plasmid library [M. J. R. Stark and J. S. Milner, *Yeast* **5**, 35 (1989)]. Two positive clones were isolated and were sequenced on both strands by standard dideoxy sequencing. The GenBank accession number is U40151.
20. An 800-bp Eco RI–Pvu II *S. cerevisiae ORC1* fragment containing the CDC-NTP domain was used as a probe on Southern blots of *S. pombe* genomic DNA under reduced hybridization stringency. Hybridizing fragments were detected that were distinct from fragments detected when a DNA fragment from *cdc18*$^{+}$ was used as a probe. The same probe was used to screen an ordered genomic cosmid library [T. Mizukami *et al.*, *Cell* **73**, 121 (1993)]. Cosmid DNA was isolated from positive clones, and hybridizing fragments were subcloned and sequenced on both strands with the use of dideoxy sequencing. A 2.0-kb Eco RI fragment was used as a probe to screen a cDNA library [J. Field *et al.*, *Mol. Cell Biol.* **8**, 2159 (1988)]. Positive clones were plaque-purified, and DNA was isolated and sequenced on both strands. The GenBank accession number is U40378.
21. Reverse transcription of human 293 total RNA was performed with degenerate primer PO1PCR7 (5′-[C/T]TCIGGIAG[A/G]TCCATIGT[A/G]TT-3′). The resulting cDNA was used as a template in PCR reactions containing primers PO1PCR7 and PO1PCR3 (5′-GT[C/G/T]CG[C/G/T][C/A]TIGA[C/T]GA[A/G]CTI- G A-3′). Products of the correct predicted size were analyzed by sequencing. Several clones encoded amino acids similar to those of *S. cerevisiae* ORC1. To identify a larger fragment, an internal exact primer (PO1PCR14, 5′- GGACCTTCTGTGGACTCACAAA-3′) was used in 3′ rapid amplification of cDNA ends (RACE) reactions containing 3′ anchor primer and 3′ adaptor primer (Gibco-BRL). Sequence analysis of amplified products revealed fragments with high homology to *S. cerevisiae ORC1*. One RACE product was used as a probe to screen a cDNA library derived from a human teratocarcinoma cell line [J. Skowronski, T. G. Fanning, M. F. Singer, *Mol. Cell. Biol.* **8**, 1385 (1988)]. Positive plaques were purified and sequenced on both strands with the use of a semiautomated DNA sequencer (Applied Biosystems). The GenBank accession number is U40152. A comparison of the human cDNA against the National Center for Biotechnology Information (NCBI) databases was done with the BLAST algorithm [S. F. Altschul *et al.*, *J. Mol. Biol.* **215**, 403 (1990)].
22. An *S. pombe orc1* knockout construct was made by digestion of pKG9, which contains a 6.0-kb Pst I genomic fragment isolated from the *S. pombe* cosmid library screen, with Not I and Sma I and filling in with Klenow to destroy the polylinker Spe I site. The plasmid was then digested with Spe I, and a 1.4-kb Spe I fragment (encoding two-thirds of the open reading frame) was replaced with the 1.8-kb *ura4*$^{+}$ gene. This construct was digested with Pst I and used to transform diploid strain SP826 (*h*$^{+N}$/*h*$^{+N}$ *ade6-210/ade6-216 leu1-32/leu1-32 ura4-D18/ura4-D18*) by standard methods [S. Moreno, A. Klar, P. Nurse, *Methods Enzymol.* **194**, 796 (1991)]. Stable Ura^{+} transformants analyzed by Southern blot carried one copy of the *orc1* disruption and one copy of the wild-type gene. After conversion of the *h*$^{+N}$ allele to *h*^{90}, tetrads from these diploids were dissected and found to segregate into two viable and two nonviable spores.
23. K. Gavin, M. Hidaka, B. Stillman, unpublished data.
24. For insertion of a HA-HIS6 epitope cassette immediately preceding the initiator methionine, a Bam HI site was introduced by PCR into pKG04, which contains a full-length *S. pombe orc1*$^{+}$ cDNA cloned into the Not I site of pSK- (Stratagene). In addition, an Nde I site was introduced at the initiator methionine to facilitate cloning into pREP1 [K. Maundrell, *Gene* **123**, 127 (1993)]. The resulting amino acid sequence at the NH_2-terminus is MGSYPYDYAHHHHHHG-SPRRKSLR . . . (*33*). The resulting plasmid (pKG14) was digested with Not I, blunt-ended with Klenow polymerase, then digested with Nde I. The vector pREP1 was digested with Sma I and Nde I and the tagged *orc1* fragment was cloned into the Nde I–Sma I sites to make pKG15.
25. Strain YB0243 (*h*^{90}/*h*$^{+N}$ *orc1::ura4/orc1*$^{+}$ *ade6-210/ade6-216 leu1-32/leu1-32 ura4-D18/ura4-D18)* was transformed with pKG15 or pKG26 {a 6-kb Hind III genomic fragment in pWH5 isolated from an *S. pombe* genomic library [M. Caligiuri and D. Beach, *Cell* **72**, 607 (1993)]} and sporulated on selective media. Colonies were streaked onto pombe minimal media containing adenine with or without thiamine.
26. K.A. Gavin and B. Stillman, unpublished data.
27. A 344-bp partial cDNA from *A. thaliana* encoding amino acids similar to those of a region of *S. cerevisiae* Orc2p was identified in the NCBI dbEST database (no. 1443). Primers based on the DNA sequence in the database (5′-CTTGCCCGGCTTTCTTCTTG-3′ and 5′-GTGTGAACAAGAAGAAAGC- CG-3′) were used to PCR-amplify a DNA fragment from an *Arabidopsis* cDNA library [D. Weigel, J. Alvarez, D. R. Smyth, M. F. Yanofsky, E. M. Meyerowitz, *Cell* **69**, 843 (1992)]. Sequences flanking this on both sides were amplified with nested PCR primers complementary to each end of the fragment, and with a primer complementary to vector sequences. The primers complementary to the 5′ end are 5′-CACTCTTGAC-CATCGGAGTT-3′ and 5′-TGCTTCCGGCTCGTAT-GTTGTGTG-3′. The primers complementary to the 3′ end are 5′-CTTGCCCGGCTTTCTTCTTG-3′ and 5′-AGTCAAGTGACTTTAAA- CTC-3′. Finally, the 5′ end of the cDNA fragment was isolated by 5′-RACE with the use of two oligonucleotides complementary to the most 5′ end of the cDNAs (5′-AGCCACACCT-GAGCTCAAAGACCCACTTTG-3′ and 5′-GCCCC-ACCCAATTCTTTTGCCAAGAAGTAG-3′). The combined clones comprised the entire *A. thaliana* cDNA. The DNA sequence was determined for all fragments on both strands. The GenBank accession number is U40269.
28. A 446-bp partial cDNA from *C. elegans* encoding amino acids related to the *S. cerevisiae* Orc2p was identified in the NCBI dbEST database (no. 16625). A partial cDNA fragment was amplified by nested PCR with the use of DNA from a *C. elegans* λZAP cDNA library (gift of R. Barstead) and oligonucleotides complementary to the dbEST cDNA sequence (5′-ACGAGGGTGAAATTGATTGC-3′, 5′-TTGTGA-ATTGACGGCAAGAG-3′ and 5′-CTACAGTTGAT-CACATTTAC-3′). The amplified product was used as a probe to screen the λZAP cDNA library. A full-length cDNA was isolated and sequenced on both strands with the use of standard dideoxy sequencing. In addition, a genomic DNA fragment containing the partial cDNA sequence identified in the dbEST database was identified in the NCBI EMBL database (no. Z36949). Computer analysis of this DNA predicts an open reading frame encoding an Orc2-related protein consisting of nine exons spanning nucleotides 1596

to 5784. However, our results from the *CeORC2* cDNA sequence revealed that the authentic protein begins at position 2618 and the first exon ends at 2793 and excludes the first five predicted exons. Before termination at 5784, the DNA from 5781 to 5882 is spliced out and the last exon terminates at 5998. The GenBank accession number is U40270.

29. A 340-bp partial cDNA encoding amino acids related to *S. cerevisiae ORC2* was isolated by RT-PCR reaction using human HeLa cell mRNA. First-strand cDNA was synthesized from 2 μg of mRNA primed with oligo(dT). The resulting cDNA was used as a template for PCR with degenerate primers based on amino acids conserved among ScORC2, AtORC2, and CeORC2 (5′-[C/T]A[C/T]GGIGTIGGITCIAA[A/G][A/C]G-3′ and 5′-ATGTGGTCIA[C/T]IG[A/T]IGCIA[C/T]IA-3′). An amplified fragment was sequenced and found to be related to *S. cerevisiae ORC2*. This combination of primers also amplified a DNA fragment from *K. lactis* that was sequenced and found to be related to *S. cerevisiae ORC2* (*23*). The human DNA fragment was used as a probe to isolate a full-length human cDNA [see (*21*)]. Positive clones were plaque-purified and sequenced on both strands. The GenBank accession number is U40268.
30. M. Gossen, D. T. S. Pak, S. Hansen, J. Acharya, M. Botchan, *Science* **270**, 1674 (1995); A. E. Ehrenhofer-Murray, M. Gossen, D. T. S. Pak, M. R. Botchan, J. Rine, *ibid.*, p. 167.
31. pKG28 contains the full-length cDNA encoding HsORC1 under the control of the cytomegalovirus promoter fused to a single copy of the T7 epitope in pCGT [A. Wilson, M. G. Peterson, W. Herr, *Genes Dev.* **9**, 2445 (1995)]. Xba I and Bam HI sites were introduced into the cDNA encoding HsORC1 by PCR immediately upstream of the in-frame initiator methionine and immediately downstream of the stop codon, respectively. The entire coding region except for the 3′-most 120 nucleotides was replaced with a native DNA fragment, and sequencing of the 120 bp at the 3′ end revealed no mutations generated during PCR. Human 293 cells were transfected or mock-transfected with 10 μg of pKG28 by electroporation and were harvested 24 hours later. Whole-cell extracts were prepared and used for immunoprecipitation with rabbit antibody to amino acids 1 to 296 of HsOrc2p fused to GST. Immune complexes were collected on protein A–Sepharose (Pharmacia), washed, and separated by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were transferred to nitrocellulose, and immunoblot analysis was done with a monoclonal antibody to the T7 epitope (Novagen). The ECL system was used for detection (Amersham).
32. A. Kornberg and T. A. Baker, *DNA Replication* (Freeman, New York, 1991), vol. 2.
33. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
34. We thank M. Hengartner, N. Kaplan, R. Martienssen, L. Rodgers, and J. Skowronski for providing various cDNA and genomic libraries; D. Casso and M. Caligiuri for *S. pombe* strains and vectors; A. Stenlund and M. Berg for help with transfections; M. Ockler for help with artwork; M. Weinreich and P. Kaufman for helpful discussions; and P. Kaufman for reading of the manuscript. Supported by the Japan Society for the Promotion of Science Postdoctoral Fellowship for Research Abroad to M.H. and by NIH grant CA13106 to B.S.

6 October 1995; accepted 13 November 1995

Separation of Origin Recognition Complex Functions by Cross-Species Complementation

Ann E. Ehrenhofer-Murray, Manfred Gossen, Daniel T. S. Pak, Michael R. Botchan, Jasper Rine*

Transcriptional silencing at the *HMR***a** locus of *Saccharomyces cerevisiae* requires the function of the origin recognition complex (ORC), the replication initiator of yeast. Expression of a *Drosophila melanogaster Orc2* complementary DNA in the yeast *orc2-1* strain, which is defective for replication and silencing, complemented the silencing defect but not the replication defect; this result indicated that the replication and silencing functions of ORC were separable. The *orc2-1* mutation mapped to the region of greatest homology between the *Drosophila* and yeast proteins. The silent state mediated by *DmOrc2* was epigenetic; it was propagated during mitotic divisions in a relatively stable way, whereas the nonsilent state was metastable. In contrast, the silent state was erased during meiosis.

Position effects are influences on the expression of a gene as a function of its location within the genome. Classical position effects include the inactivation of X chromosomes in female mammals (*1*) and position effect variegation in *Drosophila* (*2*). The inheritance of position effects in dividing cells is epigenetic, that is, genetically identical cells can display different phenotypes. At present, the mechanism of establishing and maintaining epigenetic states of gene expression is unknown. In *S. cerevisiae*, a form of position effect occurs at the cryptic mating-type loci *HMR* and *HML* (*3*). These loci serve as donors of mating-type information in mating-type interconversion. Although the genes at *HMR* and *HML* are identical to the expressed sequences at the mating-type locus *MAT*, they are kept silent throughout the cell cycle. Silencing at *HMR* and *HML* requires the combined action of nearby regulatory sequences, called silencers, and multiple proteins, at least some of which function at the silencers. Interestingly, some silencers also act as chromosomal origins of replication (*4*). Altering the autonomously replicating sequence (ARS) consensus sites of the *HMR*-E silencer results in a loss of replication initiation at the silencer and a concomitant loss of silencing at *HMR* (*5*). Moreover, mutations in two subunits of ORC decrease silencing at *HMR* and reduce replication initiation at *HMR*-E (*6*, *7*). These observations have led to the hypothesis of some link between replication and silencing. We have challenged this hypothesis by investigating the cross-species complementation of ORC functions. If a homolog is capable of complementing both functions, the association between the two processes must be ancient and conserved. In contrast, if heterologous ORC proteins complement only the silencing function or only the replication function, then the two roles of ORC are experimentally separable and may have been separate in evolution.

To test the cross-species complementation of ORC functions, we determined whether the *Drosophila Orc2* gene (*DmOrc2*) (*8*) was able to complement the *orc2-1* mutation in the yeast *ORC2* gene, which encodes the second largest subunit of ORC (*9*, *10*). A complementary DNA (cDNA) encoding DmOrc2 was placed under the control of the strong constitutive yeast glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter and the phosphoglycerate kinase terminator (PGK) (*11*). The recombinant gene was inserted into an integrative yeast–*Escherichia coli* shuttle vector and then integrated into the yeast genome (*12*). To test whether the protein encoded by *DmOrc2* was expressed in yeast, we prepared whole-cell extracts from strains that either did or did not contain the *DmOrc2* overexpression construct and analyzed them by immunoblotting with an antibody to DmOrc2 (Fig. 1). Immunoreactive DmOrc2 protein was present in the extracts from strains containing the GPD-*DmOrc2*-PGK hybrid gene, and it was not detected in strains lacking the construct. The electrophoretic mobility of the DmOrc2 protein expressed in yeast was consistent with its predicted molecular mass of 68.5 kD.

We next determined whether *DmOrc2* was capable of complementing the defects of the *orc2-1* mutation. By cloning and sequencing the *orc2-1* allele, we found that this mutation resulted in a change of a proline to a leucine residue at position 603 of the Orc2 protein, a residue that was conserved between *Drosophila* and yeast. The *orc2-1* mutation affects both the replication initiation function and the silencing function of the Orc2 protein (*6*). The rep-

A. E. Ehrenhofer-Murray and J. Rine, Division of Genetics, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.
M. Gossen, D. T. S. Pak, M. R. Botchan, Division of Biochemistry and Molecular Biology, Department of Molecular and Cell Biology, 401 Barker Hall, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: jasper@mendel.berkeley.edu