MOLECULAR AND CELLULAR BIOLOGY, June 1995, p. 3119–3128 Vol. 15, No. 6
0270-7306/95/$04.00+0

Rpa4, a Homolog of the 34-Kilodalton Subunit of the Replication Protein A Complex

KYLIE F. KESHAV, CLARK CHEN, AND ANINDYA DUTTA*

Department of Pathology, Division of Molecular Oncology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Received 14 December 1994/Returned for modification 4 January 1995/Accepted 1 March 1995

Replication protein A (RPA) is a complex of three polypeptides of 70, 34, and 13 kDa isolated from diverse eukaryotes. The complex is a single-stranded DNA-binding protein essential for simian virus 40-based DNA replication in vitro and for viability in the yeast *Saccharomyces cerevisiae*. We have identified a new 30-kDa human protein which interacts with the 70- and 13-kDa subunits of RPA, with a yeast two-hybrid/interaction trap method. This protein, Rpa4, has 47% identity with Rpa2, the 34-kDa subunit of RPA. Rpa4 associates with the 70- and 13-kDa subunits to form a trimeric complex capable of binding to single-stranded DNA. Rpa4 is preferentially expressed in placental and colon mucosa tissues. In the placenta, Rpa4 is more abundant than the 70-kDa Rpa1 subunit and is not associated with either Rpa1 or with any other single-stranded DNA-binding protein. In proliferating cells in culture, Rpa4 is considerably less abundant than Rpa1 and Rpa2. Northern (RNA) blot analysis suggests that there are alternatively processed forms of the *RPA4* mRNA, and Southern blot analysis indicates that beside *RPA4* there may be other members of the *RPA2* gene family.

Proteins that bind to single-stranded DNA (ssDNA) are essential for DNA replication and for some of the recombination and repair processes. ssDNA-binding proteins (SSBs) have been isolated from many organisms including *Escherichia coli*, bacteriophages, yeasts, and humans. They all share a common ability to remove secondary structure from ssDNA. The human SSB, known as human replication protein A (RPA), has been purified (22, 58, 60). It is a trimeric complex not only in human cells but also in other mammalian species (2, 47) and lower eukaryotes such as *Saccharomyces cerevisiae* (4, 5, 30, 31), *Xenopus laevis* (1), *Drosophila melanogaster* (44), and *Crithidia fasciculata* (6). The three peptides of human RPA have molecular masses of 70,000 (Rpa1), 34,000 (Rpa2), and 13,000 Da (Rpa3). All three polypeptides have been cloned and their encoding sequences reported (20, 21, 55). RPA binds tightly and nonspecifically to ssDNA (35) and is involved in DNA replication (22, 59, 60), repair synthesis (10, 11), and recombination (31, 45).

RPA was originally purified as a cellular component essential for DNA replication from a simian virus 40 (SV40) origin of DNA replication (22, 38, 59, 60). This replication reaction has made it possible to identify and characterize seven cellular proteins which, in combination with T antigen, are sufficient to reconstitute bidirectional, origin-dependent DNA replication (53, 57): RPA, RFC (37, 51, 54), PCNA (49–51), DNA polymerase α-primase complex (46), topoisomerases I and II (32, 63), and DNA polymerase δ (36). RPA and these host proteins are probably also involved in chromosomal DNA replication, because similar proteins have been identified in *S. cerevisiae*, in which mutations in the corresponding genes have indicated that the proteins are essential for viability and in some cases have a cell cycle block indicating a requirement for the protein in S phase (5, 23).

RPA binds tightly to ssDNA through its 70-kDa Rpa1 subunit (22, 59, 60, 62). This activity allows RPA to cooperate with T antigen to unwind an origin of DNA replication (61). Specific protein-protein interactions do not appear to play a role at this stage in vitro, as Rpa1 alone or any of a number of SSBs can substitute for RPA (33, 56). However, specific protein-protein interactions between RPA and DNA polymerase α-primase or interactions between T antigen and RPA are essential for the assembly of the initiation complex (13, 43, 53, 57). Rpa1 alone is unable to interact with T antigen (13), and heterologous SSBs cannot substitute for RPA in DNA replication (9, 20). Therefore, the 34- and 13-kDa subunits of RPA are likely to be crucial for mediating protein-protein interactions essential for DNA replication. Protein-protein interactions may also be important for the function of RPA in excision repair (10, 11, 39) and homologous recombination (31, 45). For example, the recombination complex is composed of several proteins, including human homologous protein 1 and RPA, but surprisingly, one of the subunits of RPA, Rpa2, is not detected in the recombination complex (45).

The activity of RPA may also be regulated by protein-protein interactions. The tumor suppressor protein p53 interacts with and inhibits the ability of RPA to bind ssDNA (15, 16), an interaction which may be important for controlling the onset of S phase by sequestering RPA and preventing it from binding to the replication origin. RPA is also selectively bound by the acidic activation domains of several transcription factors (27, 40), and this mechanism has been proposed to underlie the activation of DNA replication by transcription factors and the requirement for auxiliary replication-promoting sequences near the origin of DNA replication (25, 28).

Several posttranslational modifications of the RPA subunits have been observed. Both Rpa1 and Rpa2 are poly(ADP) ribosylated (18). Rpa2 is phosphorylated during the S phase of the cell cycle in both human cells and *S. cerevisiae* (12). After exposure to ionizing radiation (41) or UV light (8), human Rpa2 is phosphorylated in the G_1 phase of the cell cycle. It is not known which kinases are phosphorylating Rpa2 in vivo. In vitro, Rpa2 can be phosphorylated by cyclin B-cdc2 kinase, by cyclin A-cdk2 kinase (17, 24), and by several other DNA-dependent kinases (19, 48). Perhaps this phosphorylation is a regulatory mechanism essential for modulating RPA function

* Corresponding author. Mailing address: Department of Pathology, Division of Molecular Oncology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis St., Boston, MA 02115. Phone: (617) 278-0468. Fax: (617) 732-7449.

coordinately during the cell cycle (14, 17) and after DNA damage, and again points to the essential role that Rpa2 may play in modulating the activity of RPA.

In this paper we describe the cloning of a cDNA from a HeLa library that encodes a protein we call Rpa4. Surprisingly, this protein is about 50% identical to Rpa2 and is the first report of what may be a family of Rpa2 proteins. Selective expression of different members of the Rpa2/Rpa4 family of proteins may regulate protein-protein interactions or regulatory posttranslational modifications and thereby affect DNA replication, repair, or recombination.

MATERIALS AND METHODS

Interaction trap. An interaction trap selection for Rpa1 interactors was performed essentially as previously described (26, 64). The LexA-Rpa1 bait (16) was introduced into strain EGY48::pSH18-34 and transformed with an acid activation-tagged HeLa cDNA library (courtesy of Roger Brent). Library plasmids from transformants that showed galactose-dependent growth on Leu^- media and galactose-dependent blue color on X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) medium were rescued and sorted into several classes after restriction mapping. One representative from each class was partially sequenced on an ABI 373A automated sequencer. Primary positives were tested for target specificity by retransformation into EGY48 derivatives expressing a panel of different baits. Basal activation by the baits was determined by measuring β-galactosidase activity in cultures not expressing any acid activation-tagged proteins and was compared with the enzyme activity observed when acid-tagged proteins were expressed.

The RPA baits were made by cloning appropriate restriction fragments into pEG202 to generate LexA fusion proteins. A plasmid expressing LexA-Rpa1 was made by cloning the *Nco*I-*Sal*I fragment from pRPA70 (20) into the *Nco*I-*Xho*I sites of pEG202(b8). pEG202(b8) is derived from pEG202 by *Bam*HI digestion, Klenow treatment followed by blunt-end ligation. The resulting 4-bp insertion at the *Bam*HI site causes a frameshift so the ATG in the *Nco*I recognition sequence is in frame with the *lexA* protein. LexA-Rpa1-Nterm contained the amino-terminal 173 amino acids of Rpa1 but lacked the putative zinc finger motif (20). This was made by *Xho*I digestion of the plasmid expressing LexA-Rpa1 (*Xho*I cuts once in the Rpa1 open reading frame and once in the downstream polylinker) and subsequent self-ligation. LexA-Rpa2 and LexA-Rpa3 were made by generating the appropriate fragments by PCR so the *lexA* reading frame was in frame with the *RPA2* and *RPA3* reading frames. *RPA2* was cloned into the *Bam*HI site of pEG202, and *RPA3* was cloned into the *Eco*RI and *Xho*I sites of pEG202. LexA-Rpa2 Nterm (containing residues 1 to 102 of Rpa2) was created by deleting an *Nco*I fragment from the plasmid encoding LexA-Rpa2 (*Nco*I cuts once in the *RPA2* reading frame and once in the downstream polylinker). LexA-Rpa2-Cterm contained residues 102 to 270 and was created by cloning the *Nco*I fragment from the plasmid encoding LexA-Rpa2 into the *Nco*I site of pEG202(b8), allowing the fusion of LexA with the C-terminal part of Rpa2. LexA-Rpa2-C2term encodes amino acids 29 to 270 of Rpa2. A PCR-generated fragment encoding these amino acids was cloned into the *Eco*RI and *Bam*HI sites of pEG202.

Analysis of cDNAs encoding Rpa4. The cDNA encoding Rpa4 was excised from pJG4-5 with *Xho*I and *Eco*RI endonucleases and inserted into the *Xho*I and *Eco*RI sites of pBluescript SK^+ and pBluescript KS^+ (Stratagene), yielding pSK^+-*RPA4* (530-1565) and pKS^+-*RPA4* (530-1565), respectively. Unidirectional nested deletion sets were constructed, and the complete sequence of the cDNA was determined by manually sequencing both strands of overlapping deletions by the dideoxynucleotide chain termination method with a Sequenase kit (United States Biochemicals). The full-length Rpa4 cDNA was obtained by PCR from human placenta 5′ rapid amplification of cDNA ends (5′-RACE)-ready cDNA (Clontech) according to the manufacturer's directions. After PCR amplification, the products, ranging in size from 100 to 600 bases, were subcloned into pCRII (Invitrogen) and sequenced. We sequenced multiple independently obtained clones to ensure that there were no discrepancies introduced by PCR. Homology searches were done with the BLAST program on the NCBI blast network server.

Protein purification and antibody production. A PCR-generated fragment encoding residues 43 to 261 of Rpa4 was introduced into the *Bam*HI site of pQE12 (Qiagen), yielding pQE12-*RPA4*. This expression vector directs the synthesis of Rpa4 with a tag of six histidine residues at the N terminus. The HIS fusion protein was expressed in *E. coli* M15pREP15 and purified under denaturing conditions by affinity chromatography on nickel-nitrilotriacetic acid agarose (Qiagen) essentially as described by the manufacturer. Purified 6XHis-Rpa4 was dialyzed, and rabbit polyclonal antibodies were generated.

The entire coding sequence of Rpa2 was inserted into pRSETB (Invitrogen) as a PCR-generated *Bam*HI-*Asp* 718 fragment. The resultant 6XHis-Rpa2 protein (Rpa2 with a six-histidine residue tag at the N terminus) was overexpressed in BL21 and affinity purified on nickel-nitrilotriacetic acid agarose under denaturing conditions.

RPA was purified from suspension cultures of 293 cells as previously described (54).

Cell culture and transient expression. Colo 320 (colon carinoma) cells were maintained in RPMI 1640 medium supplemented with heat-inactiviated 10% fetal calf serum, 100 U of penicillin per ml, and 100 μg of streptomycin per ml at 37°C in 5% CO_2. Monolayer cultures of HeLa, 293T, and 293 cells were cultured under the same conditions in Dulbecco's modified Eagle's medium containing 10% calf serum and penicillin-streptomycin. 293T cells express SV40 T antigen.

A *Bam*HI fragment encoding full-length Rpa4 (405 to 1190) was generated by PCR from pKS^+-*RPA4* with 5′-cgggatccgaaggagatatatAAGATGAGTAAGAGT GG for the 5′ primer and 5′-gacggatccTCAATCAGCAGACTTAAA for the 3′ primer. This PCR fragment was inserted into the *Bam*HI site of the eukaryotic expression vector pBabe Puro, yielding pBabe Puro-*RPA4*. These constructs were transiently expressed in 293T cells. A total of 2×10^6 cells was seeded in a 6-cm-diameter dish 1 day before transfection. The transfection was performed by the calcium phosphate precipitation method (3). Cells were harvested 48 h after transfection, and protein was extracted.

Preparation of protein extracts. Cultured cells were washed twice with phosphate-buffered saline (PBS), scraped with a rubber policeman, and collected by centrifugation. Denatured extraction of protein was done by a 10-min boiling of a cell pellet from one 10-cm-diameter plate resuspended in 100 μl of 1% (wt/vol) sodium dodecyl sulfate (SDS). The lysate was adjusted to radioimmunoprecipitation assay (RIPA) buffer (10 mM Tris [pH 7.4], 5 mM EDTA, 150 mm NaCl, 10% glycerol, 0.1% SDS, 0.5% deoxycholate, 1% Triton X-100, 1% Trasylol) before being precleared by centrifugation at 4°C, at $12,000 \times g$ for 10 min. Cells were lysed under native conditions in 250 μl of RIPA buffer for 15 min on ice and then precleared as described above. Proteins were extracted from placenta, kidney, and colon tissues by homogenization on ice with a tissue homogenizer in lysis buffer (50 mM Tris [pH 7.4], 125 mM NaCl, 5 mM EDTA, 0.1% Nonidet P-40 [NP-40], 50 mM sodium fluoride, 1 mM sodium orthovanadate). After sonication on ice for five cycles (10 s on, 10 s off), the extracts were precleared by centrifugation.

Protein extracts from the following cell lines were a gift from Jon Aster: HUVEC (primary culture of human umbilical vein endothelial cells), LNCAP (prostate carcinoma), JVP1 (neuroectodermal tumor), RDES (Ewing's sarcoma), NGP and IMR32 (neuroblastoma), RH28 (rhabdomyosarcoma), HUT78 and Jurkat ($CD4^+$ T-cell leukemia), and SupT1 (ALL with *TAN-1* translocation). RIPA protein extracts were prepared from WRL68 (hepatoblastoma) and SaOs2 (osteosarcoma) cells provided by Lorene Leiter.

Protein extracts from yeast were prepared for β-galactosidase measurements (3) by the glass bead method (3).

In vitro translations and glycerol gradient sedimentation. For the creation of the functional Rpa4 for expression in vitro, the *RPA4* coding region from the *Eco*RI site at nucleotide 530 to the 3′ end was obtained from the library plasmid isolated in the interaction trap. The only part of the coding region of *RPA4* obtained by RACE PCR was upstream from the *Eco*RI site, and in this case we sequenced multiple, independently obtained clones to ensure that there were no PCR-induced discrepancies between the various clones. The complete cDNA encoding Rpa4 was constructed in pBluescript KS^+, yielding pKS^+-*RPA4* (1-1565). This plasmid was constructed by digesting pKS^+-*RPA4* (530-1565) with *Eco*RI and subsequent ligation to an *Eco*RI fragment encoding *RPA4*, bases 1 to 529. This fragment was derived from a pCRII derivative containing *RPA4* cDNA amplified from placental 5′-RACE-ready cDNA. pBluescript KS^+ constructs encoding Rpa1, Rpa2, or Rpa3 were kindly provided by T. J. Kelly. Coupled transcription-translation reactions were performed in rabbit reticulocyte lysates (Promega) with T7 polymerase. For coimmunoprecipitation and ssDNA binding experiments, the appropriate cDNAs were cotranslated.

Translation mixtures (100 μl) were loaded onto a 5-ml 10 to 40% glycerol gradient prepared in buffer A (25 mM Tris [pH 7.4], 25 mM NaCl, 1 mM EDTA, 0.01% NP-40, 1 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride). The proteins were sedimented in parallel with molecular weight protein standards in a Beckman SW55Ti rotor at 45,000 rpm for 24 h. Fractions (150 μl) were collected and analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) and fluorography.

Immunoprecipitation. Protein extracts prepared from cell lines were incubated with antibodies for 2 h at 4°C. When monoclonal antibodies were used, 1 μl of rabbit anti-mouse immunoglobulin G (Dako) was added and incubated for 30 min at 4°C. The immune complex was precipitated with 50 μl of bead slurry (1:3 [vol/vol] protein A-Sepharose–buffer). After four washes with RIPA buffer, immunoprecitated proteins were eluted from the protein A-Sepharose beads by being boiled for 5 min in 20 μl of Laemmli buffer. In vitro translation mixtures were precleared prior to immunoprecipitation by incubation with 20 μl of protein A-Sepharose bead slurry (1:3 [vol/vol] protein A-Sepharose–buffer A). The precleared translation mixture was adjusted to buffer A containing 150 mM NaCl, and the immunoprecipitations were performed as described above, except that washes were done with buffer A containing 150 mM NaCl. To dissociate protein-protein complexes before immunoprecipitation, translation mixtures (25 μl) were boiled for 5 min in 1% SDS. The boiled translation mixture was diluted to 1,000 μl with buffer A containing 150 mM NaCl.

Immunoprecipitates were analyzed by electrophoresis in SDS–15% polyacrylamide gels and transferred to nitrocellulose for immunoblotting or visualized by

TABLE 1. Interaction of Rpa4 with related proteins

Bait protein	β-Galactosidase activity[a]	
	−Rpa4	+Rpa4
LexA-Rpa1	0.5 ± 0.1	2,079.0 ± 57.0
LexA-Rpa1-Nterm	5.1 ± 0.6	397.0 ± 6.0
LexA-Rpa2	300.0 ± 93.0	ND
LexA-Rpa2-Cterm	2.1 ± 0.3	2.3 ± 0.23
LexA-Rpa2-C2term	399.1 ± 17.0	943.0 ± 93.0
LexA-Rpa2-Nterm	0.3 ± 0.02	0.5 ± 0.02
LexA-Rpa3	0.4 ± 0.01	1,791.0 ± 55.0
LexA	0.2 ± 0.01	0.3 ± 0.01

[a] β-Galactosidase levels (units per minute per milligram of protein) are averages ± standard deviations from triplicate measurements made on two independent isolates. ND, not determined.

fluorography, in the case of immunoprecipitates from in vitro-translated proteins.

Isolation of ssDNA-protein complexes. Proteins extracts or translated protein mixtures were adjusted to buffer A (25 mM Tris [pH 7.4], 25 mM NaCl, 1 mM EDTA, 0.01% NP-40, 10% glycerol, 1 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride), and incubated with ssDNA cellulose-CL6B Sepharose slurry (20 μl; 1:3 [vol/vol] ssDNA cellulose–buffer) for 2 h at 4°C. After sequential washes with buffer A and buffer A containing 200 mM NaCl, bound proteins were eluted from the ssDNA-cellulose beads by being boiled in 25 μl of Laemmli buffer for 5 min. These proteins were analyzed by electrophoresis in SDS–15% polyacrylamide gels and prepared either for immunoblotting or for autoradiography.

Western blot (immunoblot) analysis. Proteins were resolved by SDS–15% PAGE and transferred to nitrocellulose by an Owl Scientific semidry transfer apparatus. Nonspecific protein binding was blocked with BLOTTO-Tween (5% [wt/vol] nonfat dry milk, 0.1% [vol/vol] Tween 20 in PBS) for 1 h at room temperature (25°C). The blots were probed with the antibodies for 1 h either at 1:5,000 in BLOTTO-Tween for anti-Rpa4 antiserum or at 1:1,000 in 0.1% Tween-PBS for anti-RPA monoclonal antibodies (12), p70-9 (recognizes Rpa1) and p34-20 (recognizes Rpa2). After five washes with 0.1% Tween-PBS, the membranes were incubated for 30 min at room temperature with either rabbit anti-mouse or goat anti-rabbit horseradish peroxidase-conjugated immunoglobulin G (Dako; 1:1,000, diluted in 0.1% Tween-PBS). The second antibody was detected by the enhanced chemiluminescence technique (Amersham) as instructed by the manufacturer.

Southern and Northern (RNA) blotting. Polyadenylated RNA was prepared from Colo 320 cells by a single-step procedure (3) and with oligotex dT (Qiagen). mRNA (2.5 μg per lane) was separated on a 1% agarose formaldehyde gel and transferred to a nylon membrane. The blot was hybridized at room temperature in 50% formamide–5× Denhardt's solution–1 mg of denatured salmon DNA per ml–0.2% tetrasodium PP_i–1% SDS–4% dextran sulfate–3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) with random-primed [^{32}P]DNA probes corresponding to the open reading frames of *RPA2* or *RPA4* cDNA. Blots were washed in 0.1× SSC–0.5% SDS at 37°C.

Genomic DNA was isolated from HeLa cells (3), digested with *Eco*RI or *Hin*dIII, separated on a 1% agarose gel (10 μg per lane), and transferred to a nylon membrane. Hybridizations were done as described above for Northern analysis. Low-stringency hybridizations were done at 25°C with a washing at 37°C in 0.1× SSC–0.5% SDS, while high-stringency hybridizations were done at 42°C with two washes at 65°C in 0.1× SSC–0.5% SDS.

Nucleotide sequence accession number. The nucleotide sequence of *RPA4* cDNA reported in this paper will appear in the NBCI database under accession number U24186.

RESULTS

Isolation of Rpa4. To identify clones that encode proteins capable of binding to Rpa1, a HeLa cell library was screened by the interaction trap (26). When a cDNA from the library expressed a protein which interacted with the LexA-Rpa1 bait, a high level of β-galactosidase protein was expressed from pSH18-34. The latter plasmid contains a *lacZ* gene under the control of a promoter with eight *lexA* operators upstream from the TATA box. Several clones encoding proteins that reproducibly associated with Rpa1 were isolated. One of these clones (subsequently called *RPA4*) showed strong, selective interaction with Rpa1 (Table 1). When LexA-Rpa1 was present in *S. cerevisiae*, coexpression of Rpa4 from the library plasmid produced a 4,000-fold increase in the level of β-galactosidase in galactose medium in comparison with glucose medium (the library genes are under control of a galactose-inducible GAL1 promoter). The interaction required the presence of Rpa1, because no stimulation of β-galactosidase production was observed when only LexA was expressed. The C terminus of Rpa1 was required for strong interaction because a bait expressing just the N terminus produced only an 80-fold stimulation of β-galactosidase activity. A strong interaction was also detected between Rpa4 and LexA-Rpa3 bait. It was difficult to determine whether Rpa2 and Rpa4 interacted with each other because two of the Rpa2 baits, LexA-Rpa2 and LexA-Rpa2-C2term, stimulate transcriptional activity on their own. Coexpression of Rpa4 with one of these baits (LexA-Rpa2-C2term) modestly stimulated β-galactosidase activity about threefold. However, when two deletions of Rpa2 that abolished the transcriptional activity of the bait, LexA-Rpa2-Cterm and LexA-Rpa2-Nterm, were used, no significant interaction with Rpa4 was observed. A nonspecific bait, LexA fused to a newly discovered G protein, did not interact with Rpa4 (data not shown). These results provide evidence that Rpa4 interacted with high specificity with Rpa1 and Rpa3.

The 5′ end of Rpa4. Rpa4 from the interaction trap contained an insert of 1,035 nucleotides, corresponding to nucleotides 530 to 1565 in Fig. 2. Sequence analysis of this insert showed that the cDNA encoded an uninterrupted open reading frame of 219 amino acids fused in frame with the acidic transcription activation domain of the pJG4-5 vector. The clone coded for a protein which had significant identity with human Rpa2 (see Fig. 2 and 3). The homology extended up to the 5′ end of the cloned cDNA, corresponding to amino acid 43 of Rpa2 and nucleotide 530 of the *RPA4* sequence (see Fig. 2). Further, there was no methionine in Rpa4 until the methionine corresponding to amino acid 99 of Rpa2. On the basis of the extensive homology between the *RPA4* putative open reading frame upstream from this methionine with Rpa2 amino acid sequence, we hypothesized that the clone of Rpa4 was incomplete at the 5′ end and that there was an initiator methionine upstream from the N-terminal residue present in our cDNA clone. Also, antibodies raised against this Rpa4 cDNA (discussed below) recognized a larger protein than could be accounted for by a hypothetical protein beginning with the first methionine present in this cDNA.

To obtain the 5′ end of the *RPA4* cDNA, we used the rapid amplification of cDNA ends (RACE) method. A preliminary immunoblot indicated that Rpa4 was present more abundantly in placenta extracts than in HeLa cell extracts. A RACE-ready placental cDNA library, which is a single-stranded cDNA library that has an anchor oligonucleotide ligated at the 5′ end, was obtained from Clontech. After 30 cycles of PCR with the 5′ anchor oligonucleotide and with a 3′ oligonucleotide that annealed to residues 609 to 585 of Rpa4 cDNA, followed by another 30 cycles of PCR with the anchor oligonucleotide and one of two new 3′ oligonucleotides that anneal to nucleotides 567 to 541 or 572 to 546 of Rpa4 cDNA, we obtained several products, the largest being 550 bp in length (Fig. 1, lanes 1 and 2). The sequences of the cloned products (Fig. 1, lanes 4 to 10) had identical 3′ ends homologous to nucleotides 530 to 541 of the Rpa4 cDNA, indicating that they were legitimate 5′ extensions of the Rpa4 cDNA obtained in the interaction trap. The 5′ ends of the cloned products were different, presumably because of incomplete extension by the reverse transcriptase during cDNA synthesis. No evidence of alternative splicing was seen at the 5′ end of the cDNA. The sequence of these RACE products indicated that the Rpa4 protein was extended at the

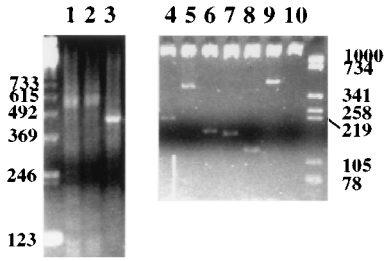

FIG. 1. Analysis of PCR products generated by the amplification of 5′ RACE-ready cDNA (Clontech) with primers specific to *RPA4* cDNA. PCR products were separated by electrophoresis on a 1% agarose gel. The secondary PCR amplification of the 5′ terminus of RPA4 cDNA was done with the anchor oligonucleotide and an oligonucleotide complementary to either nucleotides 572 to 546 (lane 1) or nucleotides 567 to 541 (lane 2). The transferrin receptor cDNA was amplified as a positive control (lane 3). Lanes 4 to 10, the amplified DNA shown in lane 1 was cloned into pCRII (Invitrogen), and insert sizes in individual clones were determined by restriction enzyme digestion.

N terminus by a further 42 amino acids, beginning with a new methionine, and the complete cDNA had an untranslated leader sequence of 404 nucleotides (Fig. 2).

Sequence analysis of the complete Rpa4 cDNA. The initiation codon of *RPA4* was assigned to nucleotides 405 to 407 and was preceded by an in-frame stop codon at nucleotides 390 to 392. This open reading frame ended with a TGA at nucleotides 1188 to 1190. The open reading frame specified a polypeptide of 261 amino acids with a predicted molecular mass of 28,880 Da. Another small open reading frame with an ATG at nucleotides 11 to 13 and a stop codon at nucleotides 122 to 124 was found upstream of *RPA4*. No significant match was found in the NBCI database with the predicted 37-amino-acid protein and, at least in rabbit reticulocyte-coupled transcription-translation reactions (see Fig. 5A), this upstream open reading frame does not interfere with *RPA4* translation.

The predicted amino acid sequence derived from the large open reading frame of the cDNA was used to search available databases with the BLAST program. This sequence was found to have homology to Rpa2, the 34-kDa subunit of the RPA complex (Fig. 3). On the basis of the homology with Rpa2, this cDNA was called *RPA4* and the protein it encodes was called Rpa4. Rpa4 and Rpa2 share 47% amino acid identity. Rpa4 also shared significant identity to Rpa2 subunits of *S. cerevisiae* and mice. The positions of sequence identity and conservative substitutions are distributed over the length of the Rpa4 protein, with a few highly conserved clusters of several residues. Like Rpa2, the N terminus of Rpa4 was unusually rich in glycine and serine. Of the first 30 amino acids, 7 (23%) were serines and 5 (17%) were glycines. As a whole, Rpa4 was not rich in glycine (5.7%) or serine (8%). Three of the four highly acidic or basic regions identified in Rpa2 (21) were conserved in Rpa4, the exception being a highly acidic stretch found between residues 95 and 123 in Rpa2. A cluster of acidic residues unique to Rpa4 was found between amino acids 186 and 195 (net charge, −5). This acidic stretch is not present in the other Rpa2 proteins.

Protein expression. A polyclonal antiserum to Rpa4 was raised in rabbits with a bacterially expressed 6XHis-Rpa4 fusion protein (Rpa4 residues 43 to 261). This antiserum specifically recognized Rpa4 and did not cross-react with a fivefold

```
5'- TCTAGTAAAAATGCATTTTTATAGAGATGTTGGGAAAGGCTTCTTGAAATTACACGTGGGACTTTTAATAGATAGGCGCTTTGACCAGCTAAGCAACAGGGCT  103

    CCCCTCGTGTGGGACTTTTAGAATGTAGCAACCACTGACACGCAGGGAAGGATTATGCGATCAGGTGAGAAGGTGGCCGACCCTGACTGGCTGGAAGCAGATG  206

    CATTCTGGTAGTTGATTGGTCCACAGGTAGCGTGACGCTTGTCACGTCCTCAGCCTCCCAGCATTCAATCGTAGCCTTTCGGACAGCTCGAAGCCCTTCTGTG  309

    GAGAGCTCGAAGCCTTCTGTGGAGAACTCAAAGCCGTCCGTGGAGCCCCAGACGAGCCAAAGCCCACCTTCTCCTCAGCCTGAGCTGTCTTGAAG ATG AGT  410
                                                                                                        M   S

    AAG AGT GGG TTT GGG AGC TAT GCG AGC ATT TCT GCT GCT GAT GGA GCG AGT GGA GGC AGT GAC CAA CTG TGT GAG AGA  488
3   K   S   G   F   G   S   Y   A   S   I   S   A   A   D   G   A   S   G   G   S   D   Q   L   C   E   R

    GAT GCA ACT CCT GCT ATT AAG ACC CAA AGA CCT AAG GTC CGA ATT CAG GAC GII GIA CCG IGI AAC GTG AAC CAG CII  566
29  D   A   T   P   A   I   K   T   Q   R   P   K   V   R   I   Q   D   V   V   P   C   N   V   N   Q   L

    CTC AGC TCT ACT GTG TTT GAC CCI GIG IIC AAG GII AGG GGA ATT ATA GTT TCC CAG GTC TCC ATC GTG GGG GTA ATC  644
55  L   S   S   T   V   F   D   P   V   F   K   V   R   G   I   I   V   S   Q   V   S   I   V   G   V   I

    AGA GGG GCA GAG AAG GCT TCA AAT CAC ATT TGT TAC AAA ATT GAT GAT ATG ACC GCG AAA CCA ATC GAG GCC CGA CAG  722
81  R   G   A   E   K   A   S   N   H   I   C   Y   K   I   D   D   M   T   A   K   P   I   E   A   R   Q

    TGG TTT GGT AGA GAG AAA GTC AAG CAG GTG ACT CCA TTG TCA GTC GGA GTA TAT GTC AAA GTG TTT GGT ATC CTC AAA  800
107 W   F   G   R   E   K   V   K   Q   V   T   P   L   S   V   G   V   Y   V   K   V   F   G   I   L   K

    TGT CCC ACG GGA ACA AAG AGC CTT GAG GTA TTG AAA ATT CAT GTC CTA GAG GAC ATG AAC GAG TTC ACC GTG CAT ATT  878
133 C   P   T   G   T   K   S   L   E   V   L   K   I   H   V   L   E   D   M   N   E   F   T   V   H   I

    CTG GAA ACG GTC AAT GCA CAC ATG ATG CTG GAT AAA GCC CGT CGT GAT ACC ACT GTA GAA AGT GTG CCT GTG TCT CCA  956
159 L   E   T   V   N   A   H   M   M   L   D   K   A   R   R   D   T   T   V   E   S   V   P   V   S   P

    TCA GAA GTG AAT GAT GCT GGG GAT AAC GAT GAG AGT CAC CGC AAT TTC ATC CAG GAC GAA GTG CTG CGT TTG ATT CAT 1034
185 S   E   V   N   D   A   G   D   N   D   E   S   H   R   N   F   I   Q   D   E   V   L   R   L   I   H

    GAG TGT CCT CAT CAG GAA GGG AAG AGC ATC CAT GAG CTC CGG GCT CAG CTC TGC GAC CTT AGC GTC AAG GCC ATC AAG 1112
211 E   C   P   H   Q   E   G   K   S   I   H   E   L   R   A   Q   L   C   D   L   S   V   K   A   I   K

    GAA GCG ATT GAT TAT CTG ACC GTT GAG GGC CAC ATC TAT CCC ACT GTG GAT CGG GAG CAT TTT AAG TCT GCT GAT TGA 1190
237 E   A   I   D   Y   L   T   V   E   G   H   I   Y   P   T   V   D   R   E   H   F   K   S   A   D   xxx

    GGCAGGGAAAACATCCTTTCATTTTTCGAAGACCCTTGCATCCAGCTGTGAGTAATTTTGACCTGTTGACTTTTTAGGAGTAGGACTAAAAAAAAAAATCTCA 1293

    AGTGGCATTCTTTGTCAACTCGCTGCTTTTCTAACTGCTTTGAACTTTTCGGATTTTCTGTATTTGAAGCTCAGAGAGAGACGGTGATGGATAAATTGACAAC 1396

    TCTGTAGGATTTACTAGCAAGCTAATGGAAACATGATTTTCGGGGAAGAAAAACTACAGAAAATGTAGAAATTTATTATTTAATTGTGTTGGAGCTTCTTTTT 1499

    CCAAAAGAAAAACTAGTTGCAGTCAGGGAGCCAGCGAAAAGACAAAAAAAAAAAAAAAAAAACACGA- 3'                                  1565
```

FIG. 2. Sequence of *RPA4* cDNA and protein. Sequence of a 1,565-nucleotide cDNA was determined by sequencing clones isolated by two different methods. Numbers on the left indicate amino acids of Rpa4. The clone isolated by screening the activation-tagged HeLa expression cDNA library encoded residues 43 to 261 and the 3′ noncoding region (corresponding to nucleotides 530 to 1565). The second set of clones, produced by amplifying placental cDNA with nested primers derived from the sequence present in the clone isolated from the yeast screen (primer sequences are underlined), encoded nucleotides 1 to 542. The underlined nucleotides indicate the sequence used to design oligonucleotides for the RACE PCR described in the legend to Fig. 1: nucleotides 609 to 585 for the primary PCR amplification, and nucleotides 567 to 541 for the secondary PCR amplification.

FIG. 3. Comparison of Rpa4 with human Rpa2 and homologs. Sequences from human Rpa2, mouse Rpa2 (mRpa2), yeast Rpa2 (yRpa2), *C. fasciculata* Rpa2 (cRpa2), and Rpa4 were aligned by the PILEUP program (Genetics Computer Group, Madison, Wis.). Residues that form the consensus are capitalized, and residues shared by Rpa4 and one or more members of the set are stippled. Numbers refer to amino acid positions in the Rpa4 sequence.

excess of Rpa2 (Fig. 4A). However, weak cross-reactivity was seen when a vast excess of Rpa2 (100 to 200 ng) was used. An Rpa2 monoclonal antibody (p34-20) did not recognize Rpa4 (12). Human RPA purified from 293 cells appeared to contain only Rpa2 and no Rpa4.

The anti-Rpa4 antiserum was used to determine the expression of Rpa4 in different tissues and cell lines. Rpa4 was abundantly expressed in placenta and, to a lesser extent, in colon tissues (Fig. 4B) as a 30-kDa protein, in good agreement with that predicted from the amino acid sequence (28.9 kDa). Interestingly, Rpa1 and Rpa2 were poorly expressed in these tissues. In contrast to the tissues, both Rpa1 and Rpa2 are abundant in all the cell lines we screened (Fig. 4C and data not shown). It is important to note that Rpa4 was observed only in three cell lines, those derived from a colon carcinoma (Colo320), a neuroblastoma (IMR32), and a Ewing's sarcoma (RDES).

In extracts from placenta and several of the cell lines, the anti-Rpa4 antibody also detected polypeptides of 80, 55, and 45 kDa (Fig. 4B and C, and see Fig. 7B), proteins not detected by preimmune serum. We do not know yet whether these proteins are related to Rpa4 or whether they are members of a large family of Rpa2-related proteins.

Association of Rpa4 with the subunits of RPA. The significant identity of Rpa4 to Rpa2 and the results from the interaction trap suggested that Rpa4 can associate with Rpa1 and with Rpa3. This was confirmed biochemically as follows. Rpa4 was synthesized together with Rpa1 and Rpa3 in rabbit reticulocyte lysate by in vitro transcription-translation. Immunoprecipitation of the products with anti-Rpa1 antibody showed coprecipitation of Rpa4 and Rpa3, suggesting that Rpa1, Rpa4, and Rpa3 form a complex (Fig. 5A). The 65-, 45-, and 28-kDa protein bands are derived from Rpa1 cDNA, by internal initiation, premature termination, or protein degradation. Rpa2, when cotranslated with Rpa1 and Rpa3, was also precipitated with anti-Rpa1 antibody and serves as a positive control. Denaturation of the putative Rpa1-4-3 complex by boiling prevented the coprecipitation of Rpa4 by anti-Rpa1 antibody. Also, negative control antibody against the p53 protein (pAb421) did not precipitate the Rpa1-4-3 complex. Rpa4 alone is not precipitated by anti-Rpa1 antibody (Fig. 5A, lane 5; also see Fig. 7b, upper left panel, and results discussed two paragraphs below).

The association of Rpa4 with Rpa1 and Rpa3 was verified by binding the cotranslated products to ssDNA cellulose. The Rpa1-2-3 complex binds tightly to ssDNA (35), with the intrinsic DNA binding activity residing in the 70-kDa Rpa1 subunit (62). Rpa4 from the in vitro transcription-translation reaction was retained on the ssDNA cellulose matrix as efficiently as Rpa2, suggesting that the association of Rpa4 with Rpa1 does not interfere with one of the functions of Rpa1, namely, binding to ssDNA (Fig. 5B). Rpa2 behaved in a similar manner. Rpa4 alone is not bound by ssDNA (see Fig. 7b, lower left panel). Also, a matrix without ssDNA (CL6B Sepharose) did not bind to the Rpa1-2-3 or Rpa1-4-3 complex.

We confirmed that the in vitro-translated Rpa1-4-3 and Rpa1-2-3 form trimeric complexes by sedimentation analysis of the proteins on a glycerol gradient (Fig. 6). Rpa4 (and Rpa2) sedimented as a broad peak ranging in size from 34 kDa (the monomer) through 45 kDa (complexed with Rpa3) to almost 150 kDA (complexed with Rpa1 and Rpa3). The Rpa3 sedimentation followed the same profile, while Rpa1 sedimented

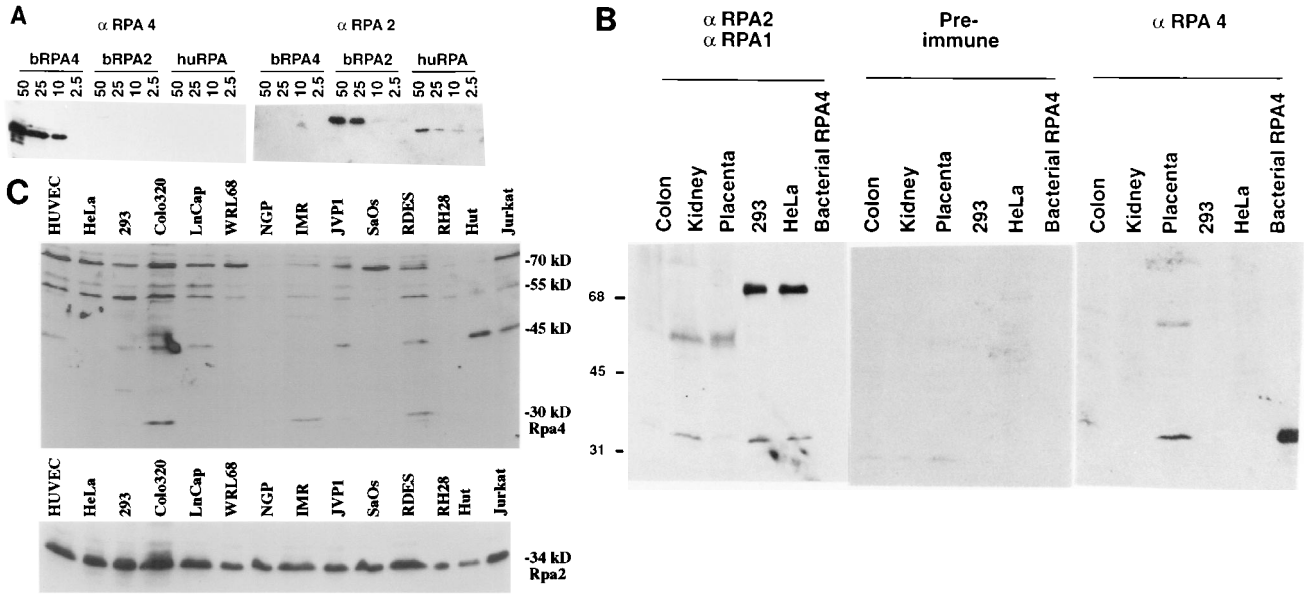

FIG. 4. Differential expression of Rpa2 and Rpa4 in cell lines and tissues. (A) Western analysis of bacterially expressed 6XHis-Rpa4 (bRPA4) and 6XHis-Rpa2 (bRPA2) and RPA purified from 293 cells (huRPA) with polyclonal antiserum raised against purified 6XHis-Rpa4 (α RPA 4) and with anti-Rpa2 monoclonal antibody (α RPA 2). The amount of protein (in nanograms) loaded in each lane is indicated at the top. (B) Protein extracts (100 μg) from tissues (placenta, kidney, and colon) or cell lines were immunoblotted with monoclonal antibodies to Rpa1 and Rpa2 (p70-9 and p34-20), preimmune Rpa4 serum, or anti-Rpa4 antiserum. The molecular masses (in kilodaltons) of the protein markers electrophoresed in parallel are indicated on the left. (C) Protein extracts (100 μg) from cell lines were immunoblotted with either anti-Rpa2 monoclonal antibody p34-20 (lower panel) or anti-Rpa4 antiserum (upper panel).

as a protein of 100 kDa and larger. This result is consistent with the formation of a Rpa1-4-3 complex of approximately 120 kDa in size. Most of Rpa1 sediments as complexes larger than 150 kDa, apparently free from Rpa2 (or Rpa4) and Rpa3. Also, when Rpa1, Rpa2, Rpa3, and Rpa4 were cotranslated together, Rpa2 and Rpa4 showed similar sedimentation profiles with no evidence of a heavier Rpa1-2-3-4 complex. Finally, Rpa4 (or Rpa2) from the lighter fractions (45 kDa) was not immunoprecipitated by anti-Rpa1 antibody (data not shown), providing further evidence that the anti-Rpa1 antibody does not cross-react with Rpa4 or Rpa2.

Since the cell lines studied expressed mostly Rpa2 with trace amounts of Rpa4, we transfected cells in culture with a plasmid expressing Rpa4 (Fig. 7A). Immunoblotting of the cell lysates showed that the transfected cells (lanes marked 2) expressed Rpa4. Immunoprecipitation of these lysates with anti-Rpa1 antibody showed that the exogenous Rpa4 was coprecipitated, indicating that when Rpa1 and Rpa4 are coexpressed they can associate with each other in vivo. Because only a small fraction of the cells are transfected and express Rpa4, while all cells express Rpa2, there is an excess of Rpa2 in the cell lysates. Cross-reaction of anti-Rpa4 antibody with the vast excess of Rpa2 probably explains the polypeptide slightly larger than Rpa4 seen in all lanes with this antibody. Immunoprecipitation of lysates containing Rpa4 with anti-Rpa2 antibody precipitated Rpa1 and Rpa3 but not Rpa4 (data not shown). Therefore, it is unlikely that Rpa1-2-4-3 or Rpa1-2-4 complexes exist.

The one instance in which Rpa4 was abundantly expressed naturally was in the placenta, but these extracts had no Rpa1. Anti-Rpa1 immunoprecipitations failed to coprecipitate the Rpa4 present in the placental extract, suggesting that Rpa4 was not physically associated with trace amounts of Rpa1 (Fig. 7B, upper left panel). ssDNA affinity experiments also showed that Rpa4 was not bound to an alternative SSB (lower left panel). The other proteins recognized by the anti-Rpa4 antiserum in placental extracts (80, 55, or 45 kDa) also did not associate with ssDNA cellulose. The absence of Rpa1 in the placental extract was unlikely to be due to failure to extract the protein, because much milder extraction conditions (25 mM NaCl, 0.01% NP-40) routinely extract Rpa1 quantitatively from cultured cells. HeLa cell extracts contain both Rpa1 and Rpa2 and have been used in parallel experiments for comparison in Fig. 7B. Rpa2 from the HeLa extracts behaved as expected from its association with Rpa1: it was coimmunoprecipitated by anti-Rpa1 antibody (upper right panel) and by the ssDNA affinity matrix (lower right panel).

In conclusion, when coexpressed with Rpa1 and Rpa3, Rpa4 can form a trimeric complex which binds to ssDNA. However, in the two tissues in which Rpa4 is expressed, Rpa1 is present

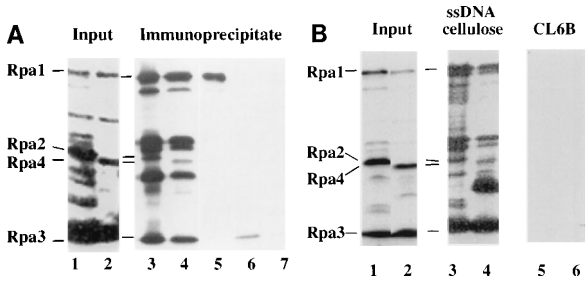

FIG. 5. Association of Rpa4 with the other subunits of RPA and with ssDNA binding activity. Proteins were detected by autoradiography. (A) Coimmunoprecipitation of Rpa4 with Rpa1 and Rpa3. Lanes 1 and 2, input; lanes 3 to 7, immunoprecipitates. Rpa1, Rpa2, and Rpa3 (lanes 1, 3, and 6, respectively) or Rpa1, Rpa4, and Rpa3 (lanes 2, 4, 5, and 7, respectively) were cotranslated and immunoprecipitated with monoclonal antibody to Rpa1, p70-9 (lanes 3 to 5), or with monoclonal antibody to p53, pAb 421 (lanes 6 and 7). Lane 5, protein-protein complexes were dissociated by being boiled in 1% SDS before immunoprecipitation. (B) Cotranslated Rpa1, Rpa2, and Rpa3 (lanes 1, 3, and 5, respectively) or Rpa1, Rpa4, and Rpa3 (lanes 2, 4, and 6, respectively) were loaded directly (lanes 1 and 2) or after binding to ssDNA cellulose (lanes 3 and 4) or Sepharose CL6B (lanes 5 and 6).

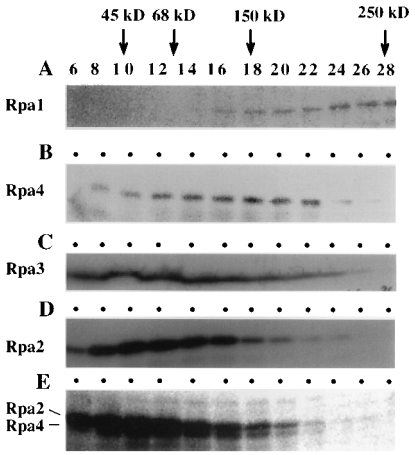

FIG. 6. Sedimentation analysis of in vitro-translated RPA complexes. The following three sets of cotranslated protein complexes were sedimented on glycerol gradients: Rpa1-2-3, Rpa1-2-3-4, and Rpa1-4-3. Gradient fraction numbers and sedimentation positions of protein standards are indicated. (A) Rpa1 sedimentation profile from Rpa1-4-3 cotranslation mixture; (B) Rpa4 sedimentation profile from Rpa1-4-3 cotranslation mixture; (C) Rpa3 sedimentation profile from Rpa1-4-3 cotranslation mixture; (D) Rpa2 sedimentation profile from Rpa1-2-3 cotranslation mixture; (E) Rpa2 and Rpa4 sedimentation profile from Rpa1-2-3-4 cotranslation mixture. Rpa1 and Rpa3 sedimentation profiles from Rpa1-2-3 and Rpa1-2-3-4 cotranslation mixtures were the same as those in panels A and C.

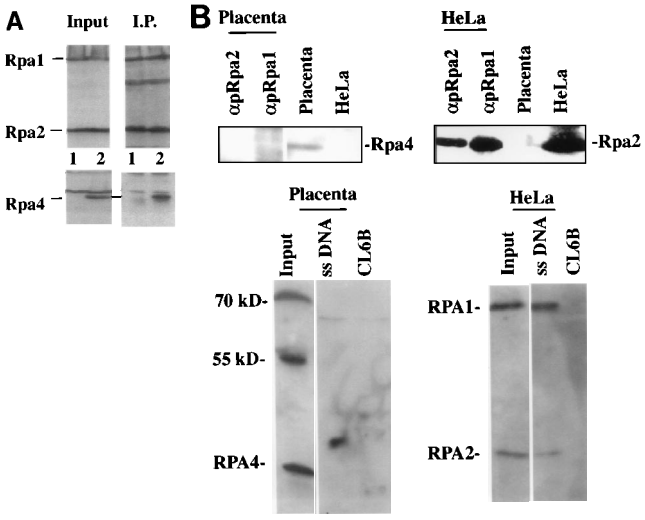

FIG. 7. Association of Rpa4 with Rpa1 in vivo. The indicated proteins are detected by immunoblotting with the appropriate antibodies. (A) 293T cells transfected with control vector pBabe Puro (lanes 1) or with pBabe Puro-*RPA4* (lanes 2). Cell lysates were either run directly on an SDS-polyacrylamide gel (Input) or after immunoprecipitation with anti-Rpa1 antibody p70-9 (I.P.). Upper plates: immunoblots with anti-Rpa1 and anti-Rpa2 antibodies. Lower plates: immunoblots with anti-Rpa4 antibody. The polypeptide slightly larger than Rpa4 is probably due to cross-reaction of the anti-Rpa4 antibody with the large excess of endogenous Rpa2. (B) Upper left, immunoblot with anti-Rpa4 antibody of placenta extract immunoprecipitated with anti-Rpa2 antibody (αRpa2; p34-20), anti-Rpa1 (p70-9), and nonimmunoprecipitated placenta and HeLa cell extracts; upper right, immunoblot with anti-Rpa2 antibody (p34-20) of HeLa cell extract immunoprecipitated with anti-Rpa2 (p34-20), anti-Rpa1 (p70-9), and nonimmunoprecipitated placenta and HeLa cell extracts; lower left, immunoblot with anti-Rpa4 antibody of placenta extract (input) and of proteins from the same extract bound to ssDNA cellulose or Sepharose CL6B beads; lower right, immunoblot with anti-Rpa1 and anti-Rpa2 antibodies of HeLa cell extract (input) and of proteins from the same extract bound to ssDNA cellulose or Sepharose CL6B beads.

in limited amounts, and Rpa4 exists free from any association with Rpa1 or any other SSB.

Other members of the Rpa2 family. The anti-Rpa4 antiserum recognized three proteins (80, 55, and 45 kDa) in addition to the 30-kDa Rpa4, so Southern blot analysis was used to determine if there were other members of the RPA2 family (Fig. 8A). The *RPA2* cDNA probe recognized additional bands at low stringency of hybridization not seen at high stringency. These additional bands (Fig. 8A, marked by dashes) did not correspond to the *RPA4*-specific fragments, so we hypothesize that these DNA fragments could encode other members of the RPA2 family. *RPA4* recognized an extra 20-kb fragment at a low stringency, but this band could correspond to *RPA2*-containing DNA.

The 80-, 55-, and 45-kDa proteins recognized by the anti-Rpa4 antibody could also have been generated by alternative splicing of the *RPA4* mRNA. Northern blot analysis of Colo 320 poly(A)$^+$ RNA with a probe to the open reading frame of *RPA4* showed three messages of 360, 500, and 700 nucleotides in addition to the expected 1,565-nucleotide-long message (Fig. 8B). Since the *RPA2* cDNA did not detect smaller messages, it is unlikely that the smaller mRNAs detected by *RPA4* cDNA were produced by nonspecific RNA degradation. Therefore, there is a possibility that alternatively processed forms of *RPA4* mRNA exist, although it is unlikely that these messages could code for the larger proteins detected with anti-Rpa4 antiserum. We have not yet detected any Rpa4-related proteins in Colo 320 cells smaller than the 30-kDa protein shown in Fig. 4C and so do not understand the significance of the small RPA4 mRNAs. Also, while Southern analysis identified *RPA2*-related DNA sequences that are different from *RPA2* itself, Northern analysis did not reveal such *RPA2*-related mRNA in Colo 320 cells. This could be because the *RPA2*-related genes were not expressed in Colo 320 cells or produce an mRNA of the same size as Rpa2.

In conclusion, Southern analysis suggests the presence of additional DNA sequences related to *RPA2*, and Northern analysis suggests the presence of alternatively processed messages derived from the *RPA4* gene.

DISCUSSION

We used the interaction trap to isolate Rpa4, a protein that interacts specifically with Rpa1. The *RPA4* cDNA spans 1,565 nucleotides containing an open reading frame that encodes a 261-amino-acid protein. The predicted molecular mass of the encoded protein is 28,880 Da, in good agreement with the molecular mass of 30,000 Da measured by SDS-PAGE. The Rpa4 protein sequence is extremely similar to that of human Rpa2. The function of Rpa2 is unknown, although we do know that it is essential for DNA replication, as antibodies directed against Rpa2 inhibit DNA replication in vitro (34) and the Rpa1 subunit alone can support DNA unwinding but not DNA replication in vitro. The derived amino acid sequence of Rpa4 diverges significantly from those of other Rpa2 proteins in two regions. An acidic region of 29 residues is present in human (amino acids 95 to 123), mouse, and yeast Rpa2 but not in Rpa4. A second region of dissimilarity is a highly acidic region present only in Rpa4 (amino acids 186 to 195). Acidic stretches are characteristic of transcriptional activator proteins (42). In the interaction trap LexA-Rpa2 was able to activate transcription, whereas Rpa2 baits containing deletions of the acidic residue region no longer activated transcription (Table 1). It is possible that the acidic regions are important for protein-protein interactions and the absence of the acidic stretch 95 to 123 in Rpa4 could be indicative of association of Rpa4 and Rpa2

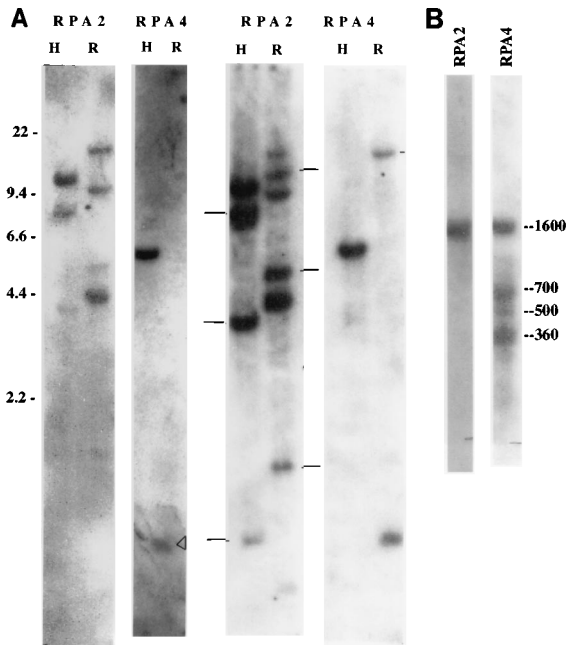

FIG. 8. *RPA2* and *RPA4* genomic and mRNA analysis. (A) Southern blot analysis of genomic DNA digested with *Eco*RI (R) or *Hin*dIII (H). DNA size (in kilobases) is indicated on the left. The first two panels are high-stringency hybridizations, and the next two are low-stringency hybridizations. DNAs visible at low stringency but not at high stringency are labeled with a dash. The *RPA4*-specific DNA fragment detected in the *Eco*RI digest is marked by an open triangle. (B) Northern analysis of polyadenylated RNA from Colo 320 cells. Probes encoding the open reading frames of either *RPA2* or *RPA4* were used in both panel A and panel B. Sizes of RNAs (in bases) are indicated on the right.

with different proteins (other than Rpa1 and Rpa3). The anti-Rpa2 antibody p34-20 significantly inhibits DNA replication, suggesting that p34-20 recognizes a part of Rpa2 which is important for DNA replication, although steric hindrance by the antibody cannot be ruled out. Because p34-20 antibody does not recognize Rpa4, this could be further evidence that Rpa4 is significantly different from Rpa2 in a domain that is important for DNA replication and contains the epitope for p34-20.

Rpa2 from both yeast and human cells is phosphorylated in a cell cycle-dependent manner (12). Several peptides of Rpa2 have been shown to be phosphorylated in vivo (17, 41). No phosphorylation was observed when Rpa2 was introduced into mouse 3T3 cells with serines 23 and 29 mutated to alanines (17). However, these serine residues are not important for SV40 replication in vitro (29). Serine 29 has been identified as the site of phosphorylation of Rpa2 by cdc2 kinase (unpublished data). Serine 29 of Rpa2 is replaced by a threonine in Rpa4, which is followed immediately by proline with several basic residues further downstream, making it a good substrate for cdc2 kinase. Phosphorylation of Rpa2 produced differences in mobility of Rpa2 in SDS-PAGE. As yet we do not have any evidence that Rpa4 is phosphorylated.

The large stretches of residues conserved between Rpa4 and Rpa2 might indicate functions conserved between the two proteins. One such conserved function is the ability of both Rpa4 and Rpa2 to associate with the ssDNA-binding Rpa1 subunit and the 13-kDa Rpa3 subunit. The interaction trap showed that Rpa4 interacted strongly with Rpa1 and Rpa3. Rpa2 also associates with Rpa1 and Rpa3 in the interaction trap (unpublished data). It has been suggested that Rpa2 fails to interact with Rpa1 to form a high-affinity RPA complex unless it is already complexed with Rpa3 (29, 52). In contrast, in the interaction trap, human Rpa4 (or Rpa2) associates with Rpa1 in the absence of human Rpa3. There could be two explanations for this apparent discrepancy. The likely explanation is that the weak association between Rpa1 and Rpa2 in the absence of Rpa3 (29, 52) also holds true for the Rpa1 and Rpa4 interaction but nonetheless is sufficiently strong to generate a positive signal in the interaction trap. A less likely explanation is that yeast Rpa3 substitutes for human Rpa3, associates with human Rpa4, and allows it to interact with the Rpa1 bait. Assuming that Rpa4 and Rpa2 interact with Rpa1 and Rpa3 in the same manner, the conserved domains between Rpa2 and Rpa4 may indicate domains necessary for these interactions.

ssDNA affinity and Rpa1 immunoprecipitation experiments suggested that Rpa4 present in the placenta was free from Rpa1 or any other SSB. We think that this is probably a reflection of the growth state of the cells in the placenta, where Rpa1 was undetectable, rather than a specific difference between Rpa2 and Rpa4. First, Rpa4 associated with Rpa1 and Rpa3 when cotranslated and also when transfected into cell lines. Second, in some circumstances Rpa2 has also been detected free from Rpa1. For instance, not all of the Rpa2 in 293 cells is associated with Rpa1, and this free Rpa2 is not phosphorylated (unpublished observations). Immunofluorescence analysis of induced myotubes has shown that Rpa1 and Rpa2 do not always colocalize during the cell cycle (7). Finally, Rpa2 is not associated with Rpa1 and Rpa3 in recombination complexes (45).

Our results suggest that Rpa4 is expressed in a tissue-specific manner, particularly in cells in quiescence. The placenta and colon mucosa extracts were the only source which had much higher levels of Rpa4 than Rpa2 (and Rpa1). Colon mucosa does have proliferating epithelial cells, but they are outnumbered by nonproliferating fibroblasts and stromal cells which constitute the bulk of the mucosa and submucosa. The placenta is also highly enriched for nonproliferating cells, particularly a syncytial cell, the syncytiotrophoblast. Therefore, the expression of Rpa4 in preference to Rpa2 may be indicative of tissue differentiation and entry of cells into quiescence. In some of the cell lines in culture (all actively proliferating), Rpa4 was expressed but at a significantly lower level than Rpa2 and Rpa1. It is interesting that Colo 320, derived from a colon cancer, expressed more Rpa4 than most of the other cell lines, concordant with the higher level of the protein in colon mucosa extracts. Therefore, the capacity to express Rpa4 may also be regulated by tissue differentiation.

Is it possible that Rpa4 is the second of a family of Rpa2-like proteins? The existence of polypeptides of 80, 55, and 45 kDa reacting with anti-Rpa4 antibody supports this hypothesis. Like Rpa4 in the placenta, these other polypeptides are not associated with Rpa1 or any other SSB. The low-stringency Southern analysis with the *RPA2* cDNA supports the possibility that at least one other member of the family which is related to *RPA2* but is not identical to *RPA2* or *RPA4* exists. Also, the *RPA4* message appears to be processed to small mRNAs coding for smaller Rpa4-related proteins, although we have not detected such proteins yet. Future experiments will be directed at clarifying whether such a family of Rpa2 proteins exist and whether there are functional differences among Rpa2, Rpa4, and other members of the putative family in DNA replication, repair, and recombination.

ACKNOWLEDGMENTS

K.F.K. was supported by an institutional training grant from the NIH, and A.D. was supported by grants from the American Cancer

Society (JFRA474) and the U.S. Armed Forces Medical Research Command (DAMD17-9h-J-4064). This work was supported by grant CA60499 from the NIH.

We thank Ellen Winchester for technical support, Jon Aster for cell line extracts, and members of the Dutta laboratory for advice and discussion.

REFERENCES

1. **Adachi, Y., and U. K. Laemmli.** 1992. Identification of nuclear pre-replication centers poised for DNA synthesis in Xenopus egg extracts: immunolocalization study of replication protein A. J. Cell Biol. **119:**1–15.
2. **Atrazhev, A., S. Zhang, and F. Grosse.** 1992. Single-stranded DNA binding protein from calf thymus. Purification, properties, and stimulation of the homologous DNA-polymerase-alpha-primase complex. Eur. J. Biochem. **210:**855–865.
3. **Ausubel, F. M., R. Brent, R. Kingston, D. Moore, J. S. Seidman, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
4. **Brill, S. J., and B. Stillman.** 1989. Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. Nature (London) **342:** 92–95.
5. **Brill, S. J., and B. Stillman.** 1991. Replication factor-A from *Saccharomyces cerevisiae* is encoded by three essential genes coordinately expressed at S phase. Genes Dev. **5:**1589–1600.
6. **Brown, G. W., T. E. Melendy, and D. S. Ray.** 1992. Conservation of structure and function of DNA replication protein A in the trypanosomatid *Crithidia fasciculata.* Proc. Natl. Acad. Sci. USA **89:**10227–10231.
7. **Cardoso, M. C., H. Leonhardt, and B. Nadal-Ginard.** 1993. Reversal of terminal differentiation and control of DNA replication: cyclin A and Cdk2 specifically localize at subnuclear sites of DNA replication. Cell **74:**979–992.
8. **Carty, M. P., M. Zernik-Kobak, S. McGrath, and K. Dixon.** 1994. UV light-induced DNA synthesis arrest in HeLA cells is associated with changes in phosphorylation of human single-stranded DNA-binding protein. EMBO J. **13:**2114–2123.
9. **Collins, K. L., and T. J. Kelly.** 1991. Effects of T antigen and replication protein A on the initiation of DNA synthesis by DNA polymerase α-primase. Mol. Cell. Biol. **11:**2108–2115.
10. **Coverley, D., M. K. Kenny, D. P. Lane, and R. D. Wood.** 1992. A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. Nucleic Acids Res. **20:**3873–3880.
11. **Coverley, D., M. K. Kenny, M. Munn, W. D. Rupp, D. P. Lane, and R. D. Wood.** 1991. Requirement for the replication protein SSB in human DNA excision repair. Nature (London) **349:**538–541.
12. **Din, S., S. Brill, M. P. Fairman, and B. Stillman.** 1990. Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells. Genes Dev. **4:**968–977.
13. **Dornreiter, I., L. F. Erdile, I. U. Gilbert, D. von Winkler, T. J. Kelly, and E. Fanning.** 1992. Interaction of DNA polymerase α-primase with cellular replication protein A and SV40 T antigen. EMBO J. **11:**769–776.
14. **D'Urso, G., R. L. Marraccino, D. R. Marshak, and J. M. Roberts.** 1990. Cell cycle control of DNA replication by a homologue from human cells of the $p34^{cdc2}$ protein kinase. Science **250:**786–791.
15. **Dutta, A.** 1993. Trans-plication factors? Curr. Biol. **3:**709–712.
16. **Dutta, A., J. M. Ruppert, J. C. Aster, and W. Winchester.** 1993. Inhibition of DNA replication factor RPA by p53. Nature (London) **365:**79–82.
17. **Dutta, A., and B. Stillman.** 1992. *cdc2* family kinases phosphorylate a human cell DNA replication factor, RPA, and activate DNA replication. EMBO J. **11:**2189–2199.
18. **Eki, T., and J. Hurwitz.** 1991. Influence of poly (ADP-ribose) polymerase on the enzymatic synthesis of SV40 DNA. J. Biol. Chem. **266:**3087–3100.
19. **Elledge, S. J., R. Richman, F. L. Hall, R. T. Williams, N. Lodgson, and J. W. Harper.** 1992. CDK2 encodes a 33-kDa cyclin A-associated protein kinase and is expressed before CDC2 in the cell cycle. Proc. Natl. Acad. Sci. USA **89:**2907–2911.
20. **Erdile, L. F., W.-D. Heyer, R. Kolodner, and T. J. Kelly.** 1991. Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subunit of human replication protein A and the role of the protein in DNA replication. J. Biol. Chem. **266:**12090–12098.
21. **Erdile, L. F., M. S. Wold, and T. J. Kelly.** 1990. The primary structure of the 32-kDa subunit of human replication protein A. J. Biol. Chem. **265:**3177–3182.
22. **Fairman, M. P., and B. Stillman.** 1988. Cellular factors required for multiple stages of SV40 replication *in vitro*. EMBO J. **7:**1211–1218.
23. **Fien, K., and B. Stillman.** 1992. Identification of replication factor C from *Saccharomyces cerevisiae*: a component of the leading-strand DNA replication complex. Mol. Cell. Biol. **12:**155–163.
24. **Fotedar, R., and J. M. Roberts.** 1992. Cell cycle regulated phosphorylation of RPA-32 occurs within the replication initiation complex. EMBO J. **11:**2177–2187.
25. **Guo, Z. S., and M. L. Depamphilis.** 1992. Specific transcription factors stimulate simian virus 40 and polyomavirus origins of DNA replication. Mol. Cell. Biol. **12:**2514–2524.
26. **Gyuris, J., E. Golemis, H. Chertkov, and R. Brent.** 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. Cell **75:**791–803.
27. **He, Z., B. T. Brinton, J. Greenblatt, J. A. Hassell, and C. J. Ingles.** 1993. The transactivator proteins VP16 and gal4 bind replication factor A. Cell **73:** 1223–1232.
28. **Heintz, N. H.** 1992. Transcription factors and the control of DNA replication. Curr. Opin. Cell Biol. **4:**459–467.
29. **Henricksen, L. A., C. B. Umbricht, and M. S. Wold.** 1994. Recombinant replication protein A: expression, complex formation and functional characterization. J. Biol. Chem. **269:**11121–11132.
30. **Heyer, W. D., and R. D. Kolodner.** 1989. Purification and characterization of a protein from *Saccharomyces cerevisiae* that binds tightly to single-stranded DNA and stimulates a cognate strand exchange protein. Biochemistry **28:** 2856–2862.
31. **Heyer, W.-D., M. R. S. Rao, L. F. Erdile, T. J. Kelly, and R. D. Kolodner.** 1990. An essential *Saccharomyces cerevisiae* single-stranded DNA binding protein is homologous to the large subunit of human RP-A. EMBO J. **9:**2321–2329.
32. **Ishimi, Y., K. Sugasawa, F. Hanaoka, T. Eki, and J. Hurwitz.** 1992. Topoisomerase II plays an essential role as a swivelase in the late stage of SV40 chromosome replication *in vitro*. J. Biol. Chem. **267:**462–466.
33. **Kenny, M. K., S.-H. Lee, and J. Hurwitz.** 1989. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases α and δ. Proc. Natl. Acad. Sci. USA **86:**9757–9761.
34. **Kenny, M. K., U. Schlegel, H. Furneaux, and J. Hurwitz.** 1990. The role of human single stranded DNA binding protein and its individual subunits in simian virus 40 DNA replication. J. Biol. Chem. **265:**7693–7700.
35. **Kim, C., R. O. Snyder, and M. S. Wold.** 1992. Binding properties of replication protein A from human and yeast cells. Mol. Cell. Biol. **12:**3050–3059.
36. **Lee, S. H., T. Eki, and J. Hurwitz.** 1989. Synthesis of DNA containing the simian 40 origin of replication by the combined action of DNA polymerases alpha and delta. Proc. Natl. Acad. Sci. USA **86:**7361–7365.
37. **Lee, S.-H., and J. Hurwitz.** 1990. Mechanism of elongation of primed DNA by DNA polymerase delta, proliferating cell nuclear antigen, and activator 1. Proc. Natl. Acad. Sci. USA **87:**5672–5676.
38. **Li, J. J., and T. J. Kelly.** 1984. Simian virus 40 DNA replication *in vitro*. Proc. Natl. Acad. Sci. USA **81:**6973–6977.
39. **Li, L., S. J. Elledge, C. A. Peterson, E. S. Bales, and R. J. Legerski.** 1994. Specific association between the human DNA repair proteins XPA and ERCC1. Proc. Natl. Acad. Sci. USA **91:**5012–5016.
40. **Li, R., and M. R. Botchan.** 1993. The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate *in vitro* BPV-1 DNA replication. Cell **73:**1207–1221.
41. **Liu, V. F., and D. T. Weaver.** 1993. The ionizing radiation-induced replication protein A phosphorylation response differs between ataxia telangiectasia and normal human cells. Mol. Cell. Biol. **13:**7222–7231.
42. **Ma, J., and M. Ptashne.** 1987. A new class of yeast transcriptional activators. Cell **51:**113–119.
43. **Melendy, T., and B. Stillman.** 1993. An interaction between replication protein A and SV40 T antigen appears essential for primosome assembly during SV40 DNA replication. J. Biol. Chem. **268:**3389–3395.
44. **Mitsis, P. G., S. C. Kowalczykowski, and I. R. Lehman.** 1993. A single-stranded DNA binding protein from *Drosophila melanogaster*: characterization of the heterotrimeric protein and its interaction with single-stranded DNA. Biochemistry **32:**5257–5266.
45. **Moore, S. P., L. Erdile, T. Kelly, and R. Fishel.** 1991. The human homologous pairing protein HPP-1 is specifically stimulated by the cognate single-stranded binding protein hRP-A. Proc. Natl. Acad. Sci. USA **88:**9067–9071.
46. **Murakami, Y., C. R. Wobbe, L. Weissbach, F. B. Dean, and J. Hurwitz.** 1986. Role of DNA polymerase α and DNA primase in simian virus 40 DNA replication *in vitro*. Proc. Natl. Acad. Sci. USA **83:**2869–2873.
47. **Nakagawa, M., S. Tsukada, T. Soma, S. Yoshifumi, S. Miyake, A. Iwamatsu, and H. Sugiyama.** 1991. cDNA cloning of the murine 30 kDa protein homologous to the 32 kDa subunit of human replication protein A. Nucleic Acids Res. **19:**4292.
48. **Pan, Z. Q., A. Amin, and J. Hurwitz.** 1993. Characterization of the *in vitro* reconstituted cyclin A or B1-dependent cdk2 and cdc2 kinase activities. J. Biol. Chem. **268:**20443–20451.
49. **Prelich, G., M. Kostura, D. R. Marshak, M. B. Mathews, and B. Stillman.** 1987. The cell cycle-regulated proliferating cell nuclear antigen is required for SV40 DNA replication *in vitro*. Nature (London) **326:**471–475.
50. **Prelich, G., and B. Stillman.** 1988. Coordinated leading and lagging strand synthesis during SV40 DNA replication *in vitro* requires PCNA. Cell **53:**117–126.
51. **Prelich, G., C.-K. Tan, M. Kostura, M. B. Mathews, A. G. So, K. M. Downey, and B. Stillman.** 1987. Functional identity of proliferating cell nuclear antigen and a DNA polymerase δ auxiliary protein. Nature (London) **326:**517–520.
52. **Stigger, E., F. B. Dean, J. Hurwitz, and S. H. Lee.** 1994. Reconstitution of

functional human single-stranded DNA-binding protein from individual subunits expressed by recombinant baculoviruses. Proc. Natl. Acad. Sci. USA **91:**579–583.
53. **Tsurimoto, T., T. Melendy, and B. Stillman.** 1990. Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV 40 DNA replication origin. Nature (London) **346:**534–539.
54. **Tsurimoto, T., and B. Stillman.** 1989. Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication *in vitro*. EMBO J. **8:**3883–3889.
55. **Umbricht, C. B., L. F. Erdile, E. W. Jabs, and T. J. Kelly.** 1993. Cloning, overexpression and genomic mapping of the 14-kDa subunit of human replication protein A. J. Biol. Chem. **268:**6131–6138.
56. **Virshup, D. M., and T. J. Kelly.** 1989. Purification of replication factor C, a cellular protein involved in the initial stages of simian virus 40 DNA replication *in vitro*. Proc. Natl. Acad. Sci. USA **86:**3584–3588.
57. **Weinberg, D. H., K. L. Collins, P. Simancek, A. Russo, M. S. Wold, D. M. Virshup, and T. J. Kelly.** 1990. Reconstitution of simian virus 40 DNA replication with purified proteins. Proc. Natl. Acad. Sci. USA **87:**8692–8696.
58. **Wobbe, C. R., F. Dean, L. Weissbach, and J. Hurwitz.** 1985. *In vitro* replication of duplex circular DNA containing the simian virus 40 origin site. Proc. Natl. Acad. Sci. USA **82:**1834–1838.
59. **Wobbe, C. R., L. Weissbach, J. A. Borowiec, F. B. Dean, Y. Murakami, P. Bullock, and J. Hurwitz.** 1987. Replication of SV40 origin containing DNA *in vitro* with purified proteins. Proc. Natl. Acad. Sci. USA **84:**1834–1838.
60. **Wold, M. S., and T. Kelly.** 1988. Purification and characterization of replication protein A, a cellular protein required for *in vitro* replication of simian virus 40 DNA. Proc. Natl. Acad. Sci. USA **85:**2523–2527.
61. **Wold, M. S., J. J. Li, and T. J. Kelly.** 1987. Initiation of simian virus 40 DNA replication *in vitro*: large tumor antigen and origin-dependent unwinding of the template. Proc. Natl. Acad. Sci. USA **84:**3643–3637.
62. **Wold, M. S., D. W. Weinberg, D. M. Virshup, J. J. Li, and T. J. Kelly.** 1989. Identification of cellular proteins required for SV40 DNA replication. J. Biol. Chem. **264:**2801–2809.
63. **Yang, L., M. S. Wold, J. J. Li, T. J. Kelly, and L. F. Lui.** 1987. Roles of DNA topoisomerases in simian virus 40 DNA replication. Proc. Natl. Acad. Sci. USA **84:**950–954.
64. **Zervos, A. S., J. Gyuris, and R. Brent.** 1993. Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. Cell **72:**223–232.