MOLECULAR AND CELLULAR BIOLOGY, Dec. 1992, p. 5620–5631
0270-7306/92/125620-12$02.00/0

Vol. 12, No. 12

# Molecular Cloning of Cellular Genes Encoding Retinoblastoma-Associated Proteins: Identification of a Gene with Properties of the Transcription Factor E2F

BEI SHAN, XUELIANG ZHU, PHANG-LANG CHEN, TIM DURFEE, YANZHU YANG, DAVE SHARP, AND WEN-HWA LEE*

*Center for Molecular Medicine/Institute of Biotechnology, University of Texas Health Science Center at San Antonio, 15355 Lambda Drive, San Antonio, Texas 78245*

Received 31 July 1992/Returned for modification 4 September 1992/Accepted 30 September 1992

**The retinoblastoma protein interacts with a number of cellular proteins to form complexes which are probably crucial for its normal physiological function. To identify these proteins, we isolated nine distinct clones by direct screening of cDNA expression libraries using purified RB protein as a probe. One of these clones, Ap12, is expressed predominantly at the $G_1$-S boundary and in the S phase of the cell cycle. The nucleotide sequence of Ap12 has features characteristic of transcription factors. The C-terminal region binds to unphosphorylated RB in regions similar to those to which T antigen binds and contains a transactivation domain. A region containing a potential leucine zipper flanked by basic residues is able to bind an E2F recognition sequence specifically. Expression of Ap12 in mammalian cells significantly enhances E2F-dependent transcriptional activity. These results suggest that Ap12 encodes a protein with properties known to be characteristic of transcription factor E2F.**

The retinoblastoma gene (RB), the first tumor suppressor gene identified, encodes a nuclear phosphoprotein which is ubiquitously expressed in vertebrates (19, 20, 42, 44). Mutations of this gene which lead to inactivation of its normal function have been found not only in 100% of retinoblastomas but also in many other adult cancers, including small-cell lung carcinoma (25, 69), osteosarcoma (64), bladder carcinoma (28), prostate carcinoma (4), and breast cancer (41). Reconstitution of a variety of RB-deficient tumor cells with wild-type RB leads to suppression of their neoplastic phenotypes, including the ability to form tumors in nude mice (5, 10, 21, 30, 62, 63). These results provide direct evidence that RB protein is an authentic tumor suppressor.

Although the underlying mechanism by which RB suppresses tumor formation is not clear, several important clues have been gathered in recent studies. RB appears to function in the $G_0$-$G_1$ phase of the cell cycle, as substantiated by several findings: (i) phosphorylation of RB, presumably by members of the cyclin-dependent kinase (Cdk) family (45, 47), fluctuates with the cell cycle (7, 11, 13); (ii) the unphosphorylated form of RB is present predominantly in the $G_0$-$G_1$ stage (11, 13); (iii) microinjection of unphosphorylated RB into cells at early $G_1$ inhibits their progression into the S phase (22). These data suggest that RB serves as a critical regulator of entry into the cell cycle and that its inactivation in normal cells leads to deregulated growth.

How RB functions is the subject of intense inquiry. Two known biochemical properties of the RB protein have been described; one is its intrinsic DNA-binding activity, which was mapped to its C-terminal 300 amino acid residues (44, 66); another is its ability to interact with several oncoproteins of DNA tumor viruses (12, 16, 67). This interaction was mapped to two discontinuous regions at amino acids 394 to 571 and 649 to 773, designated as the T-binding domains (29, 33). Interestingly, mutations of RB protein in tumors were frequently located in these same regions (3). These results imply that the T-binding domains of RB protein are functionally important and the interaction of RB with these oncoproteins may have profound biological significance. The identification of cellular proteins that mimic the binding of T to RB revealed a potentially complicated network. Several proteins, including c-*myc* (53), Rb-p1, Rb-p2 (14), and 8 to 10 other proteins (31, 36, 43), have been shown to bind to RB in vitro. Recently, several groups reported that RB interacts with transcription factor E2F (1, 2, 9). Although the gene for E2F has not been cloned and its identity is based solely on the ability to recognize and bind to a specific DNA sequence, these studies suggested that tight complexes involving RB, E2F, and other proteins exist in cells (8, 16, 59). To identify the cellular proteins with which RB interacts, we used an "RB sandwich" method to screen two λgt11 expression libraries by using a purified RB (amino acids 379 to 928) that includes both T-binding domains and the entire C-terminal region as a probe. Nine distinct cDNA clones were isolated, one of which, Ap12, has characteristics similar to that of transcription factor E2F. Similar results were described in two recent reports (26, 35).

## MATERIALS AND METHODS

**Isolation of cellular genes that encode RB-associated proteins.** Two cDNA libraries were constructed from poly(A)$^+$ RNAs isolated from HeLa and Saos2 cells by previously described methods (55). The double-stranded cDNAs were size fractionated by using Sepharose CL-4B chromatography and then ligated to λgt11 arms. The sizes of the in vitro-packaged libraries were $2.0 \times 10^7$ recombinants for HeLa cells and $1.5 \times 10^7$ for Saos2 cells, with an average insert size of 1.6 kb. The cDNA libraries were plated on 100 150-mm-diameter dishes at $1 \times 10^4$ to $2 \times 10^4$ recombinants per dish and incubated at 42°C until plaques just became visible (3.5 h) and then transferred to nitrocellulose filters saturated with isopropyl-β-D-thiogalactopyranoside (IPTG;

* Corresponding author.

10 mM) and left overnight at 37°C. The filters were denatured in 6 M guanidine HCl, renatured, and incubated with the RB sandwich probe in binding buffer (25 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.5], 50 mM NaCl, 5 mM $MgCl_2$, 5 mM dithiothreitol, 0.1% Nonidet P-40 [NP-40], 5% milk, 1 mg of bovine serum albumin per ml) for 4 hours at 4°C. The RB sandwich was prepared by mixing 1 μg of purified bacterially expressed p56-RB (31), 10 μl of preabsorbed polyclonal anti-RB antibody (anti-RB 0.47; 1:100 dilution), and 1 μl of an alkaline phosphatase-conjugated secondary antibody (1:1,000 dilution) per ml of binding buffer, incubated at 4°C for 2 h. The control sandwich was prepared by mixing the RB antibody and the secondary antibody and used as a control to eliminate the clones cross-reacted with the anti-RB antibody. The bound filters were then washed in TBST (20 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.05% Tween 20) five times for 3 min each time and color developed with 5-bromo-4-chloro-3-indolylphosphate nitroblue tetrazolium (BCIP/NBP; Promega, Madison, Wis.). Positive clones from the initial screening were picked and subjected to second and third rounds of screening. The clones that consistently showed positive signals with the RB sandwich but not with the RB-negative sandwich were then selected for fourth and fifth rounds of screening by plating at a low density with a control bacteriophage to ensure homogeneous isolates which gave strong positive signals over the background.

**Plasmid construction and fusion protein expression.** The cDNA inserts of RB-associated protein (RbAp) clones were subcloned into pGEM1 for sequence analysis. To express fusion RbAps in vitro, the cDNA inserts were ligated in frame into the glutathione S-transferase (Gst) fusion protein expression system (60). Expression of the Gst fusion proteins was induced with 0.2 mM IPTG, and the bacterial lysates were prepared by two rounds of freezing and thawing, followed by sonication in lysis buffer B (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 5 mM dithiothreitol, 0.2% NP-40, 1 mM phenylmethylsulfonyl fluoride, 1 μg of leupeptin per ml, 5 μg of aprotinin per ml, 1 μg of antipain per ml) and clarified by centrifugation. The bacterially expressed Gst fusions were prepared and purified by using Gst agarose beads.

**In vitro binding assay.** The binding assay was performed as follows. Beads containing about 1 to 2 μg of Gst, Gst-Ap12, or Gst-T were mixed with Molt4 cell lysates ($2 \times 10^6$ cells) or bacterial lysates expressing pETRB in 400 μl of lysis buffer B at 4°C for 60 min. The bound beads were subsequently washed five times in 1 ml of phosphate-buffered saline–0.2% NP-40, and the protein complex was boiled in sodium dodecyl sulfate (SDS) loading buffer. RB bound to the Gst fusions was analyzed by immunoblotting by using monoclonal anti-RB antibody mAb11D7 (41a), which recognizes the C-terminal region of RB protein, or mAb245, which recognizes a region containing amino acids 371 to 390.

**Construction of mutated RB proteins expressed in the bacterial pET-T7 system.** In addition to pETRbc, pETM6, and pETM9 (31), pETB2, pETSsp, and pETM8 were constructed by cloning *Aha*II-*Bam*HI fragments from pB2, pSsp, and pM8 (33) into the corresponding pET expression vector. The bacterial lysates were prepared as described in the previous section.

**Construction of Gst-RbAp12 fusion proteins.** The DNA fragments derived from RbAp12 clones were subcloned into the Gst fusion plasmids. Gst-P3 was constructed by cloning the *Eco*RI-*Sph*I fragment from the original 1.4-kb cDNA (G12) into pGEPK (9a), a derivative of pGEX-2T (60). Gst-SH5 contains the *Sma*I-*Hin*dIII fragment from clone B6, and Gst-XH9 contains the *Eco*RI-*Hin*dIII fragment of clone A6, which contains the entire coding sequence. Gst-SX4 and Gst-XX4 are derived from Gst-SH5 and Gst-XH9, respectively, but the 3′ end of the *Xho*I-*Hin*dIII fragment is deleted. Gst-SA1, Gst-BB8, and Gst-BH7 are all derived from Gst-SH5 and contain deletion mutations of the *Sal*I-*Acc*I fragment for SA1, the *Bam*HI-*Bsa*BI fragment for BB8, and the *Bsa*BI-*Hin*dIII fragment for BH7, respectively. Gst-HH1 was generated by subcloning the *Sal*I-*Hin*cII fragment of Ap12 into the Gst fusion vector.

**RNA blot analysis.** Total RNA (10 μg) extracted by the guanidine isothiocyanate-CsCl method (55) was denatured in 50% formamide–2.2 M formaldehyde–20 mM Na borate (pH 8.3) and analyzed by 1.0% agarose gel electrophoresis. The RNA was then transferred to Hybond paper (Amersham), and the RNA was immobilized by UV cross-linking. Prehybridization and hybridization were carried out in 50% formamide–5× SSPE (1× SSPE is 0.18 M NaCl, 10 mM $NaPO_4$, and 1 mM EDTA [pH 7.7])–5× Denhardt solution–1% SDS–100 μg of salmon sperm DNA per ml, and hybridization was performed in the presence of the $^{32}$P-labeled 1.4-kb RbAp12 insert DNA at 45°C for 18 h. The initial washing was done with 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at room temperature, and the final washing was done with 0.1× SSC–0.1% SDS at 65°C for 45 min.

**DNA gel mobility shift assay.** A plasmid DNA fragment containing two E2F recognition sequences (TTTCGCGC and GCGCGAAA) was used as a probe for the gel mobility shift assay and also served as a competitor. A DNA fragment containing a mutated E2F site (TTTAGCGC—GCGCTA AA) (32), which does not bind to E2F, was also used as a competitor. The assay was performed as described previously (68). The diluted Gst-Ap12 bacterial lysates were incubated with either the wild type, a mutant, or no competitor (1 μg) in binding buffer (20 mM HEPES [pH 7.6], 1 mM $MgCl_2$, 0.1 mM EGTA [ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N'*,*N'*-tetraacetic acid], 40 mM KCl, 10% glycerol)–0.1% NP-40–1 mg of salmon sperm DNA per ml at room temperature for 15 min, and the $^{32}$P-end-labeled (Klenow fill-in) probe (10 ng) was added; this was followed by another 30 min of incubation. The protein-DNA complexes were analyzed by 4% acrylamide gel electrophoresis in 0.25× TBE buffer at 4°C.

**Yeast expression vector and strain.** The expression plasmid used in *Saccharomyces cerevisiae* was based on the pAS1 vector (15a). Briefly, the plasmid contains the yeast ADH1 promoter driving expression of the yeast GAL4 DNA-binding domain, followed by a downstream polylinker. The vector also carries the 2μm origin and *TRP1* gene for maintenance and selection in *S. cerevisiae*. pAS/G12 was constructed by subcloning the *Eco*RI fragment isolated from G12 into the unique *Eco*RI site in pAS1. Similarly, pAS/B6 was built by subcloning the *Eco*RI fragment from B6 into the pAS1 *Eco*RI site. pASRb2 will be described elsewhere (15a). The *S. cerevisiae* strain used was Y153 [*MAT***a** *trp1-901 leu2-3,112 ade2-101 ura3-52*::*URA3* (*GAL1-lacZ*) *MEL* (*GAL1-lacZ*)] (14a).

**Yeast transformation and β-galactosidase assay.** Yeast transformation was carried out by using the LiOAc method as described previously (56). After transformation, cells were plated on synthetic dropout medium lacking tryptophan to select for the presence of the plasmid. Following 2 to 3 days of growth at 30°C, single colonies from each transformation were streaked onto another selective plate and allowed to grow for an additional 24 h. The colony color β-galactosidase activity assay was then performed as previ-

ously described (6), except that the nitrocellulose filters were submerged in liquid nitrogen for about 30 to 60 s to permeabilize the cells and then thawed at room temperature before being overlaid on Whatman filters saturated with LacZ–X-Gal solution (6). The color developed in about 20 min in the case of the AP12 clones. No color change was observed with the pAS/Rb2 clone alone, even after overnight exposure.

**Transient transfection assay.** Transfections were carried out with CV1 cells by conventional calcium phosphate precipitation. For GAL4-Ap12 chimeric plasmids, *Eco*RI fragments isolated from Gst-P3, Gst-XH9, and Gst-XX4 were subcloned in frame into the pSG424 expression vector containing the GAL4 DNA-binding domain driven by the simian virus 40 (SV40) early promoter (17, 54). The E1BCAT and $GAL4_5$E1BCAT reporter plasmids were described previously (46). Expression plasmid pCMVAp12Stu was constructed by cloning the *Stu*I fragment from clone A6 into the *Sma*I site of pCMV, and pCMVAp12RH contains the *Eco*RI-*Hin*dIII fragment of clone B6. Plasmid pCMVAp12ΔDS was constructed from pCMVAp12Stu by deletion of the *Dra*III-*Sal*I fragment (nucleotides 343 to 568). The cytomegalovirus constructs (5 μg) were cotransfected with 5 μg of plasmid pE2F$A_{10}$CAT (containing two E2F-binding sites), pE2FmA$_{10}$CAT (containing two mutated E2F-binding sites), or p$A_{10}$CAT (containing no E2F-binding sites) into the same number of cells ($5 \times 10^6$). A plasmid expressing the *lacZ* gene was also included in the transfection assay as an internal control. Chloramphenicol acetyltransferase (CAT) activities were measured after 48 h as described previously (23).

## RESULTS

**Identification of RbAps.** Two λgt11 cDNA expression libraries were constructed and screened by using the purified p56-RB protein (amino acids 376 to 928), which includes both T-binding domains and the entire C-terminal region (31), as the probe. We refer to this probe as an RB sandwich since it contains RB protein, rabbit anti-RB antibody 0.47 (33), and alkaline phosphatase-conjugated goat anti-rabbit immunoglobulin G (see Materials and Methods). Figure 1 contains a diagram of the sandwich screening strategy (panels A and B). Since the association of RB and SV40 T antigen is well documented (12), a λgt11 phage expressing T antigen was constructed, screened with the RB sandwich, and used as a positive control (Fig. 1D). As an example (Fig. 1C), the fusion product of one of the clones (Ap12) was readily detected by this method. One half of each filter was used for binding to the RB sandwich, and the other half was used for binding to the sandwich minus RB protein. The latter probe served as a control for background binding due to any cross-reaction of the RB antibody or goat anti-rabbit antibody with bacterial proteins. After five rounds of screening of $10^6$ recombinant phage, 12 clones emerged as genes that encode candidate RB-associated proteins. They were named RbAp1, -2, -4, -6, -8, -9, -10, -11, -12, -13, -14, and -15.

These 12 putative RbAp cDNAs were subcloned into the pGEM plasmid, and a partial sequence of 500 to 600 bp from each clone was obtained. When compared with the known gene sequences present in the GENBANK data base, RbAp1, -2, -4, -8, -10, -12, -13, -14, and -15 appeared to be novel genes that contain no significant homology to any known genes. However, three clones matched previously identified genes: RbAp6 is identical to nuclear lamin C (18, 48); RbAp9 encodes a product of human papillomavirus type 16 (HPV) E7 (52); and RbAp11 codes for the upstream

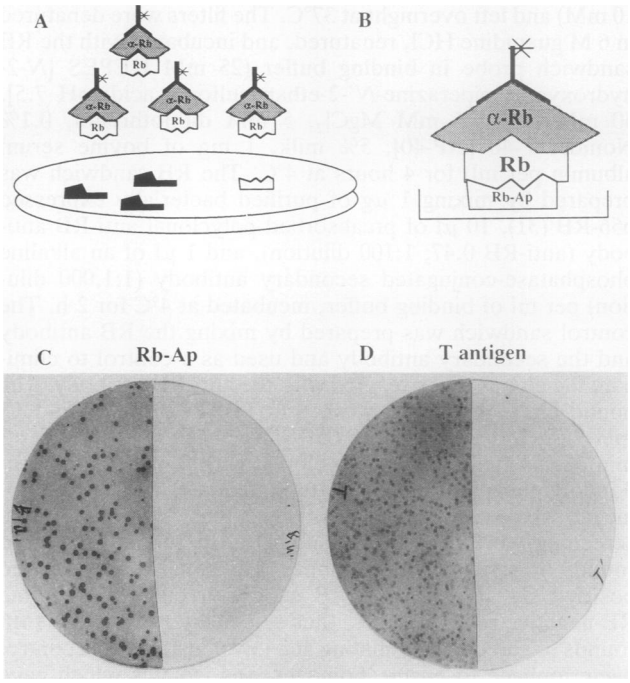



FIG. 1. RB sandwich screening. λgt11 cDNA expression libraries were plated and screened by using the RB sandwich containing purified p56-RB, an anti-RB antibody, and an alkaline phosphatase-conjugated secondary antibody. (A and B) Diagram of the RB sandwich screening procedure. (C and D) Hybridized filters with the RB sandwich (left halves of the filters) in which the positive signal indicates an RbAp-RB complex (C) or a T-antigen–RB complex (D). The right halves of the filters were probed with the sandwich without RB protein.

binding factor that binds to the ribosomal RNA gene promoter (34). Cross-hybridization and sequencing data showed that RbAp1, -10, -13, and -14 are identical. Table 1 summarizes the preliminary characterization of these cloned RbAps.

**Expression of Ap12 mRNA occurs predominantly at the $G_1$-S boundary and in the S phase of the cell cycle.** Among the clones that consistently showed strong binding during

TABLE 1. Initial characterization of RbAps[a]

| RbAp(s) | Length of cDNA (kb) | Size of mRNA (kb) | In vitro binding | Identity | Subcellular localization |
|---|---|---|---|---|---|
| 1, 10, 13, 14 | 2.8 | 7.1 | + | Novel | Nucleus |
| 2 | 1.6 | 3.6 | ND | Novel | ND |
| 4 | 1.7 | 6.7 | + | Novel | Nucleus |
| 6 | 1.5 | 2.1 | + | Lamin C | Nucleus |
| 8 | 1.8 | 6.9 | ND | Novel | ND |
| 9 | 0.5 | ND | ND | HPV-E7 | ND |
| 11 | 1.5 | 3.2 | + | UBF | Nucleus |
| 12 | 1.4 | 2.6 | + | Novel | ND |
| 15 | 1.5 | 6.5 | + | Novel | Nucleus |

[a] The size of the cDNA of each clone was determined by ethidium bromide staining of the agarose gel after digestion of the phage DNA with *Eco*RI. The sizes of mRNAs were measured by RNA blot analysis using 28S and 18S rRNAs as markers. A partial sequence of each clone was used to search the GENBANK data base to determine the identity of the clone. Nuclear localization was determined by immunostaining and cell fractionation (data not shown). ND, not determined. UBF, upstream binding factor.

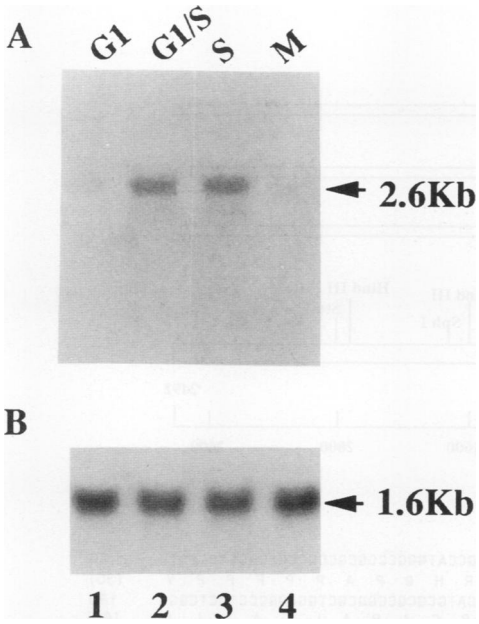



FIG. 2. Predominant expression of Ap12 at the $G_1$-S boundary and in the S phase in the cell cycle. Normal monkey kidney CV1 cells were plated in fresh Dulbecco modified Eagle medium with 10% fetal calf serum in the presence of lovastatin (50 μM) for 36 h ($G_1$) or aphidicolin (10 μg/ml) for 16 h ($G_1$-S) and then released for 4 h (S) or incubated in the presence of nacodazole (0.4 μg/ml) for another 16 h (M). Total RNA was extracted from the cells synchronized at these stages of the cell cycle. Ten micrograms of the RNA was denatured and analyzed by formaldehyde gel electrophoresis. Two sets of RNA blots were hybridized with $^{32}$P-labeled cDNAs of Ap12 (G12) (A) or the Gβ-like gene (B). Lanes: 1, early $G_1$; 2, $G_1$-S boundary; 3, S phase; 4, M phase. mRNA size (arrow) was determined by migration of 28S and 18S rRNAs, which were run in a parallel lane next to the RNA samples.

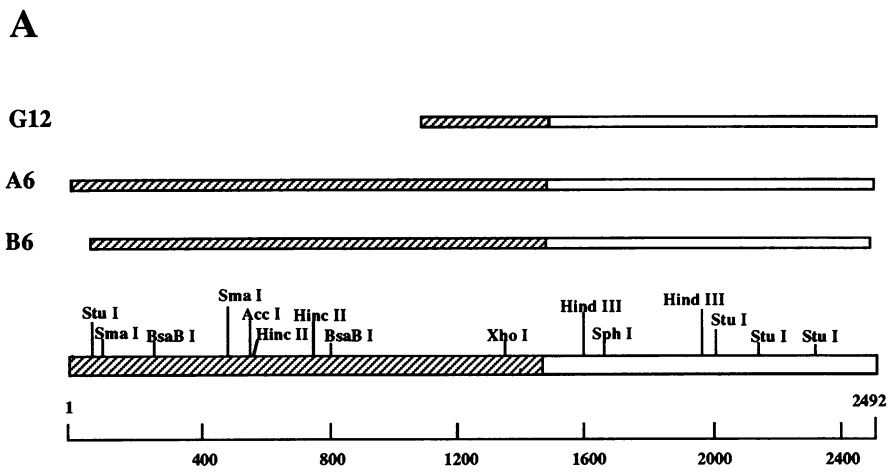
screening was HPV E7 protein, a viral oncoprotein known to associate with RB (16). This was not unexpected since one of the cDNA libraries was constructed from HeLa cells, an HPV-infected cell line. Another clone, Ap12, showing even stronger binding than E7 during screening, is a previously unidentified gene. The clone has an insert of 1.4 kb with an approximately 1.0-kb untranslated region and contains an open reading frame of 114 amino acids. We performed RNA blot analysis to determine the size of the mRNA and its pattern of expression during cell cycle progression. Normal monkey kidney CV1 cells were plated in fresh medium with 10% fetal calf serum in the presence of lovastatin for 36 h (to arrest the cells in the $G_1$ phase) (38) or aphidicolin for 16 h (to arrest the cells at the $G_1$-S boundary) and then released for 4 h (to allow the cells to enter the S phase) or incubated in the presence of nocodazole for another 16 h (to arrest the cells in the M phase) (22). Total RNA from each stage was prepared for blot analysis by using Ap12 cDNA as a probe. As shown in Fig. 2, an approximately 2.6-kb single species of mRNA was predominantly expressed at the $G_1$-S boundary and in the S phase but was undetectable in the early $G_1$ or M phase. These results are consistent with the data described by Kaelin et al. (35). As a control, the expression of a Gβ-like gene (24) was tested and verified to be constant during the cell cycle (Fig. 2), suggesting that Ap12 is involved in the regulation of cell cycle progression.

**Sequence analysis of Ap12.** The initial Ap12 cDNA clone (G12, Fig. 3) was shorter than its corresponding mRNA. Accordingly, cDNA libraries were rescreened to obtain longer clones; two of these clones, A6 and B6, are shown in Fig. 3. The longest clone, A6, contains a cDNA sequence of 2,492 nucleotides. It appears that the sequence of Ap12 is identical to those previously reported (26, 35), although the length of the 5′ sequence is slightly different. Since there is no stop codon in front of the first methioine, there are two alternatives for deducing the full length of the putative protein. The candidate protein may be initiated at the position of amino acid residue 40 (circled in Fig. 3) and contains a total of 437 amino acids, as described by others (26, 35). Alternatively, the protein may contain more than 476 amino acids and its first methionine has not been identified. We thus tentatively numbered the amino acid on the basis of the longest known open reading frame, since there is not sufficient data to distinguish between these two alternatives.

Distinctive features of the putative protein sequence include about 100 C-terminal amino acids that are very acidic and an N-terminal region dominated by proline residues. Following the proline-rich region are putative leucine repeats (39) flanked by stretches of basic amino acids, suggesting a potential DNA-binding domain. These features are indicative of several different classes of eukaryotic transcription factors (39). Furthermore, there are two potential phosphorylation sites for Cdk kinase (57) at amino acids 159 to 161 (KSP) and 346 to 349 (SPGK), which could modulate the function of this protein.

**Ap12 binds only the hypophosphorylated form of RB at regions similar to those required for binding of SV40 T antigen.** To analyze the RB-binding properties of Ap12, the original clone (G12) was expressed as a Gst fusion protein (P3) and purified by glutathione agarose chromatography. This fusion protein was tested for the ability to bind full-length RB prepared from a cellular lysate of Molt4 cells that expressed both hyper- and hypophosphorylated forms of the RB protein. A Gst–T-antigen fusion protein and Gst alone were also tested for this ability. As shown in Fig. 4A, the P3 protein bound only to the hypophosphorylated form of RB and the binding signal was very similar to that of T. Gst alone did not bind RB protein. To define the sequences of RB required for binding of Ap12, a panel of RB mutants (Fig. 4C) expressed in the bacterial pET-T7 expression system (61) were mixed with P3 beads or, in parallel, with Gst-T beads. The amount of wild-type or mutated RB protein bound to the beads was determined by Western blot (immunoblot) analysis by using monoclonal anti-RB antibody mAb245 (43). As shown in Fig. 4B and D, wild-type RB (pETRbc) and two mutant RB proteins (B2 and M9), in which deletions were outside the T-binding domains, bound to both T (panel B) and Ap12 (panel D), while other RB mutants (pETSsP, XS, M6, M8, and NM) defective in binding to T also failed to bind to Ap12. These results indicate that Ap12 and T require similar regions of the unphosphorylated form of RB to bind, suggesting that the Ap12-RB association is biologically significant.

**The C-terminal region of Ap12 is required for binding to RB.** Since the initial P3 fusion protein that contains 114 amino acids of Ap12 binds to RB, additional experiments were designed to map the region of Ap12 required for binding to RB. Four Gst-Ap12 fusion proteins with different N- or C-terminal deletions were constructed (Fig. 5B). XH9 contains the entire coding sequence of Ap12 cDNA, and SH5 (from *Sma*I to *Hin*dIII) contains the C-terminal 314 amino acids. XX4 and SX4 are derived from XH9 and SH5,




B

```
CTTTGCAGGCAGCGGCGGCCGGGGGCGGAGCGGGATCGAGCCCTCGCCGAGGCCTGCCGCCATGGGCCCGCGCCGCCGCCGCCGCCTGTC    90
 L  C  R  Q  R  R  P  G  A  E  R  D  R  A  L  A  E  A  C  R  H  G  P  A  P  P  P  P  P  V   (30)
ACCCGGGCCGCGCGGGCCGTGAGCGTCATGGCCTTGGCCGGGGCCCCTGCGGGCGGCCCATGCGCGCCGGCGCTGGAGGCCCTGCTCGGG   180
 T  R  A  A  R  A  V  S  V  M  A  L  A  G  A  P  A  G  G  P  C  A  P  A  L  E  A  L  L  G   (60)
GCCGGCGCGCTGCGGCTGCTCGACTCCTCGCAGATCGTCATCATCTCCGCCGCGCAGGACGCCAGCGCCCCGCCGGCTCCCACCGGCCCC   270
 A  G  A  L  R  L  L  D  S  S  Q  I  V  I  I  S  A  A  Q  D  A  S  A  P  P  A  P  T  G  P   (90)
GCGGCGCCCGCCGCCGGCCCCTGCGACCCTGACCTGCTGCTCTTCGCCACACCGCAGGCGCCCCGGCCCACACCCAGTGCGCCGCGGCCC   360
 A  A  P  A  A  G  P  C  D  P  D  L  L  L  F  A  T  P  Q  A  P  R  P  T  P  S  A  P  R  P   (120)
GCGCTCGGCCGCCCGCCGGTGAAGCGGAGGCTGGACCTGGAAACTGACCATCAGTACCTGGCCGAGAGCAGTGGGCCAGCTCGGGGCAGA   450
 A  L  G  R  P  P  V  K  R  R  L  D  L  E  T  D  H  Q  Y  L  A  E  S  S  G  P  A  R  G  R   (150)
GGCCGCCATCCAGGAAAAGGTGTGAAATCCCCGGGGGAGAAGTCACGCTATGAGACCTCACTGAATCTGACCACCAAGCGCTTCCTGGAG   540
 G  R  H  P  G  K  G  V  K  S  P  G  E  K  S  R  Y  E  T  S  L  N  L  T  T  K  R  F  L  E   (180)
CTGCTGAGCCACTCGGCTGACGGTGTCGTCGACCTGAACTGGGCTGCCGAGGTGCTGAAGGTGCAGAAGCGGCGCATCTATGACATCACC   630
 L  L  S  H  S  A  D  G  V  V  D  L  N  W  A  A  E  V  L  K  V  Q  K  R  R  I  Y  D  I  T   (210)
AACGTCCTTGAGGGCATCCAGCTCATTGCCAAGAAGTCCAAGAACCACATCCAGTGGCTGGGCAGCCACACCACAGTGGGCGTCGGCGGA   720
 N  V  L  E  G  I  Q  L  I  A  K  K  S  K  N  H  I  Q  W  L  G  S  H  T  T  V  G  V  G  G   (240)
CGGCTTGAGGGGTTGACCCAGGACCTCCGACAGCTGCAGGAGAGCGAGCAGCAGCTGGACCACCTGATGAATATCTGTACTACGCAGCTG   810
 R  L  E  G  L  T  Q  D  L  R  Q  L  Q  E  S  E  Q  Q  L  D  H  L  M  N  I  C  T  T  Q  L   (270)
CGCCTGCTCTCCGAGGACACTGACAGCCAGCGCCTGGCCTACGTGACGTGTCAGGACCTTCGTAGCATTGCAGACCCTGCAGAGCAGATG   900
 R  L  L  S  E  D  T  D  S  Q  R  L  A  Y  V  T  C  Q  D  L  R  S  I  A  D  P  A  E  Q  M   (300)
GTTATGGTGATCAAAGCCCCTCCTGAGACCCAGCTCCAAGCCGTGGACTCTTCGGAGAACTTTCAGATCTCCCTTAAGAGCAAACAAGGC   990
 V  M  V  I  K  A  P  P  E  T  Q  L  Q  A  V  D  S  S  E  N  F  Q  I  S  L  K  S  K  Q  G   (330)
CCGATCGATGTTTTCCTGTGCCCTGAGGAGACCGTAGGTGGGATCAGCCCTGGGAAGACCCCATCCCAGGAGGTCACTTCTGAGGAGGAG  1080
 P  I  D  V  F  L  C  P  E  E  T  V  G  G  I  S  P  G  K  T  P  S  Q  E  V  T  S  E  E  E   (360)
AACAGGGCCACTGACTCTGCCACCATAGTGTCACCACCACCATCATCTCCCCCCTCATCCCTCACCACAGATCCCAGCCAGTCTCTACTC  1170
 N  R  A  T  D  S  A  T  I  V  S  P  P  P  S  S  P  P  S  S  L  T  T  D  P  S  Q  S  L  L   (390)
AGCCTGGAGCAAGAACCGCTGTTGTCCCGGATGGGCAGCCTGCGGGCTCCCGTGGACGAGGACCGCCTGTCCCCGCTGGTGGCGGCCGAC  1260
 S  L  E  Q  E  P  L  L  S  R  M  G  S  L  R  A  P  V  D  E  D  R  L  S  P  L  V  A  A  D   (420)
TCGCTCCTGGAGCATGTGCGGGAGGACTTCTCCGGCCTCCTCCCTGAGGAGTTCATCAGCCTTTCCCCACCCCACGAGGCCCTCGACTAC  1350
 S  L  L  E  H  V  R  E  D  F  S  G  L  L  P  E  E  F  I  S  L  S  P  P  H  E  A  L  D  Y   (450)
CACTTCGGCCTCGAGGAGGGCGAGGGCATCAGAGACCTCTTCGACTGTGACTTTGGGGACCTCACCCCCCTGGATTTCTGACAGGGCTTG  1440
 H  F  G  L  E  E  G  E  G  I  R  D  L  F  D  C  D  F  G  D  L  T  P  L  D  F  .            (476)
GAGGGACCAGGGTTTCCAGAGATGCTCACCTTGTCTCTGCAGCCCTGGAGCCCCCTGTCCCTGGCCGTCCTCCCAGCCTGTTTGGAAACA  1530
TTTAATTTATACCCCTCTCCTCTGTCTCCAGAAGCTTCTAGCTCTGGGGTCTGGCTACCGCTAGGAGGCTGAGCAAGCCAGGAAGGGAAG
GAGTCTGTGTGGTGTGTATGTGCATGCAGCCTACACCCACACGTGTGTACCGGGGGTGAATGTGTGTGAGCATGTGTGTGTGCATGTACC
GGGGAATGAAGGTGAACATACACCTCTGTGTGTGCACTGCAGACACGCCCCAGTGTGTCCACATGTGTGTGCATGAGTCCATGTGTGCGC
GTGGGGGGGCTCTAACTGCACTTTCGGCCCTTTTGCTCTGGGGGTCCCACAAGGCCCAGGGCAGTGCCTGCTCCCAGAATCTGGTGCTCT
GACCAGGCCAGGTGGGGAGGCTTTGGCTGGCTGGGCGTGTAGGACGGTGAGAGCACTTCTGTCTTAAAGGTTTTTTCTGATTGAAGCTTT
AATGGAGCGTTATTTATTTATCGAGGCCTCTTTGGTGAGCCTGGGGAATCAGCAAAGGGGAGGAGGGGTGTGGGGTTGATACCCCAACTC
CCTCTACCCTTGAGCAAGGGCAGGGGTCCCTGAGCTGTTCTTCTGCCCCATACTGAAGGAACTGAGGCCTGGGTGATTTATTTATTGGGA
AAGTGAGGGAGGGAGACAGACTGACTGACAGCCATGGGTGGTCAGATGGTGGGGTGGGCCCTCTCCAGGGGGCCAGTTCAGGGCCCCAGC
TGCCCCCCAGGATGGATATGAGATGGGAGAGGTGAGTGGGGGACCTTCACTGATGTGGGCAGGAGGGGTGGTGAAGGCCTCCCCCAGCCC
AGACCCTGTGGTCCCTCCTGCAGTGTCTGAAGCGCCTGCCTCCCCACTGCTCTGCCCCACCCTCCAATCTGCACTTTGATTTGCTTCCTA
ACAGCTCTGTTCCCTCCTGCTTTGGTTTTAATAAATATTTTGATGACGTTAAAAAAAAAAAA  2492
```

FIG. 3. Restriction map and sequence of Ap12. (A) Restriction map of three cDNA clones of Ap12: G12, A6, and B6. Only restriction sites used to construct Ap12 derivatives are shown. (B) Complete sequence of clone A6, including 2,492 nucleotides and its predicted amino acid sequence of 476 amino acids (in parentheses). The first amino acid is designated as position one. The first in-frame methionine is at amino acid 40 (circled). The squares indicate potential leucine repeats. Two putative Cdk phosphorylation sites are underlined.

respectively, and contain a deletion of 21 amino acids at the C terminus. The bacterially expressed RB protein (pETRbc) was mixed with these Gst-Ap12 derivatives and analyzed by Western blotting using anti-RB antibody mAb245 as described above. As shown in Fig. 5A, XH9, SH5, and P3 bound similar amounts of RB, suggesting that the additional N-terminal sequence of Ap12 contributes little to RB binding. However, XX4 and SX4, which both have 21 amino

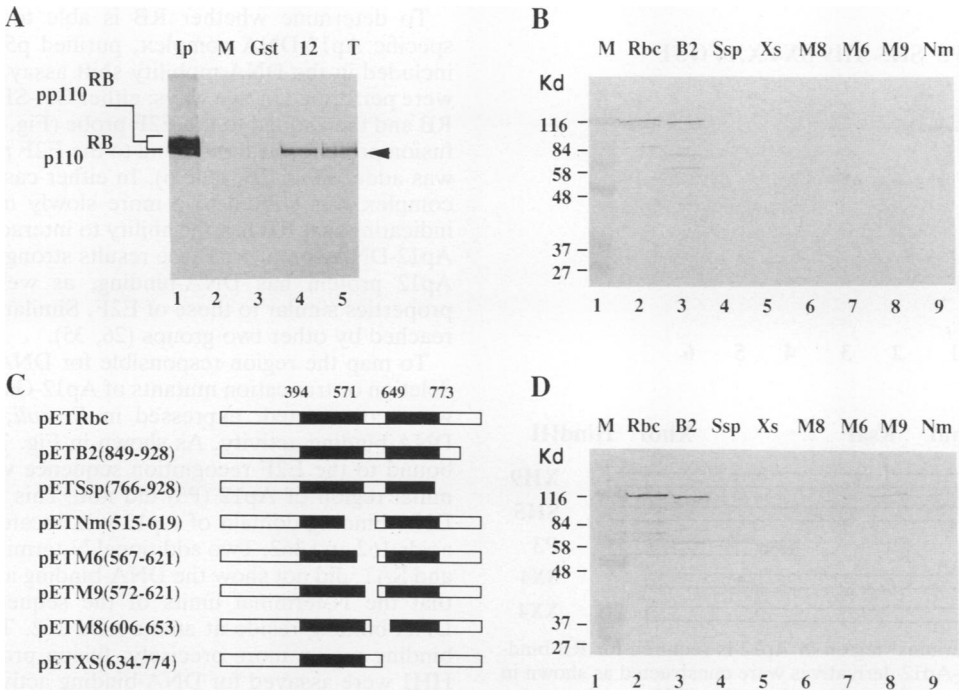



FIG. 4. Ap12 binds specifically to the hypophosphorylated form of RB at regions similar to those bound by T. (A) Ap12 and T antigen bind to the hypophosphorylated form of RB from Molt4 cells. Gst fusion proteins of Ap12 and SV40 T antigen were prepared and purified by using glutathione beads. Beads containing about 1 to 2 μg of Gst (lane 3), Gst-Ap12 (lane 4), or Gst-T (lane 5) were mixed with Molt4 cell lysates ($2 \times 10^6$ cells) in 400 μl of lysis buffer as described in Materials and Methods. RB bound to the Gst fusions was analyzed by immunoblotting using monoclonal anti-RB antibody mAb11D7, which recognizes the C-terminal region of the RB protein. Lanes: 1, total RB protein from a Molt4 cell lysate immunoprecipitated by using monoclonal anti-RB antibody mAb11D7; 2, molecular weight marker. (C) Panel of RB mutant proteins expressed in a bacterial pET-T7 expression system. The numbers in parentheses indicate the deleted amino acids of the RB protein. The T-binding domains are highlighted. (B and D) Ap12 (D) binds to RB in regions similar to those bound by T (B). The bacterial lysates containing about 300 ng of wild-type RB protein (pETRbc) or mutant RB proteins (pETB2, Ssp, Xs, M8, M6, M9, and Nm) were mixed with beads containing 1 to 2 μg of Gst-Ap12 (D) or Gst-T (B) for the binding assay as described for panel A, except that mAb245 was used. The binding patterns of Ap12 and T are very similar, although the signal that show binding of both Ap12 and T to M9 are reduced.

acids deleted from the C terminus, failed to bind RB. Taken together, these results indicate that the C-terminal region of Ap12 is required for binding to RB, which is consistent with the results previously described (26).

**The C terminus of Ap12 can function as a transcription activation domain.** Highly acidic, amphipathic alpha-helical regions commonly serve as activation domains in eukaryotic transcription factors (49). The C-terminal region of Ap12 also displays these characteristics, suggesting that it functions in an analogous manner. To test this, both clones B6 (encoding amino acids 22 to 476) and G12 (amino acids 362 to 476) were fused to the DNA-binding domain of the yeast GAL4 protein (amino acids 1 to 147) (37) present on a yeast expression vector. While this GAL4 fragment can bind specifically to its recognition site ($UAS_G$) (37), it lacks an activation domain. Therefore, the chimeric protein relies on the fused segment to provide transcriptional activation for a $UAS_G$-containing promoter. Several such fusions involving mammalian activators have been shown to be functional in *S. cerevisiae*, including GAL4-p53 (17). As shown in Fig. 6A, following transformation of a yeast strain harboring the *Escherichia coli lacZ* gene under $UAS_G$ control, both GAL4-Ap12 fusions were able to activate transcription of the reporter, as evidenced by β-galactosidase activity, whereas the GAL4-RB control was not. The GAL4 DNA-binding domain alone showed little or no β-galactosidase activity (data not shown).

To characterize the transactivation activity of Ap12 further, similar experiments were performed with mammalian CV1 cells in which the CAT reporter plasmids containing no ($G_0$E1BCAT) or five ($G_5$E1BCAT) GAL4 DNA-binding sites were cotransfected with the chimeric plasmids expressing GAL4-Ap12 fusions. As shown in Fig. 6B, either the C-terminal region of Ap12 (SGP3) or the full-length Ap12 GAL4 fusion (SGXH9) displayed strong transactivation activity when cotransfected with $G_5$E1BCAT but not when transfected with $G_0$E1BCAT, whereas the GAL4 DNA-binding domain alone (SG424) had no effect. Similar results were reported by Kaelin et al. (35). As a positive control, a strong activator, VP16, when fused to the GAL4 DNA-binding domain (SGVP16), also showed strong transactivation activity (52). These data, from both yeast and mammalian cells, indicate that the C-terminal 114 amino acids of Ap12 contain a strong transactivation domain.

Interestingly, a mutant construct containing a C-terminal 21-amino-acid deletion of Ap12 (SGXX4), which abolished the RB-binding ability, retained a significant amount of transactivation activity, suggesting that the RB-binding domain is dissociated from the transactivation activity of Ap12. Contrary to this result, Kaelin et al. reported that the C-terminal 20 amino acids were required for transactivation (35). The cause of this discrepancy is unclear, and further experiments are needed to map the precise sequences required for transactivation.

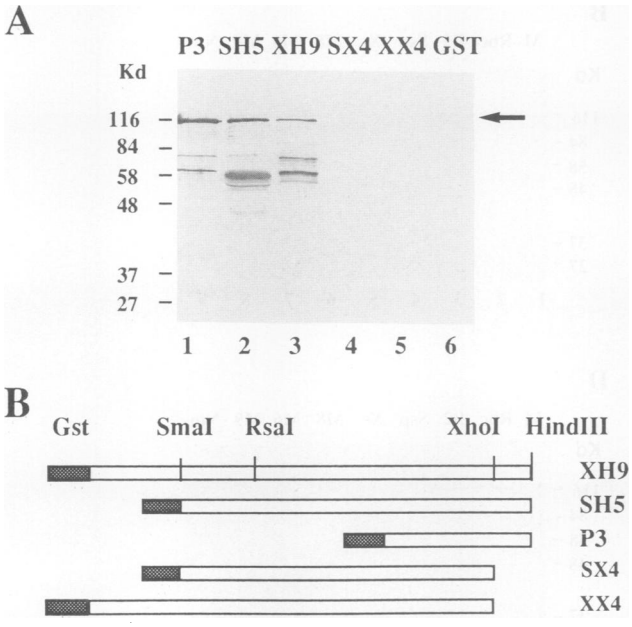



FIG. 5. The C-terminal region of Ap12 is required for RB binding. A series of Gst-Ap12 derivatives were constructed as shown in panel B. P3 contains amino acids 362 to 476, SH5 contains amino acids 162 to 476; XH9 contains the full open reading frame (1 to 476), SX4 contains amino acids 162 to 455, and XX4 contains amino acids 1 to 455. About 300 ng of bacterially expressed pETRbc (wild-type RB) was mixed with these Gst-Ap12 beads or with Gst beads which were preadjusted to contain the same amount of proteins (about 1 μg). The binding conditions were as described above, and the bound RB proteins were analyzed by Western blotting with monoclonal anti-RB antibody mAb245. The arrow indicates the position of p110-RB.

**Ap12 binds specifically to the E2F recognition sequence.** Since RB forms a complex with transcription factor E2F (1, 2, 9) and Ap12 has a potential DNA-binding domain, bacterially expressed Gst-Ap12 fusion proteins were tested for the ability to bind a DNA fragment containing two E2F recognition sites (68) by using a DNA mobility shift assay. As shown in Fig. 7A, the lysate prepared from SH5 (Fig. 7D) was able to bind the DNA fragment and the binding was specific because the complex could be effectively competed with the excess, unlabeled wild-type E2F cognate sequence but not with a mutated sequence that differs from the wild type by only one nucleotide (32). Similar results were obtained when XH9 was used in the assay (data not shown). Partially purified E2F protein from adenovirus-infected HeLa cells also specifically bound to the DNA probe as previously reported (68). Two different DNA-protein complexes were observed when purified HeLa E2F was used, compared with one complex when the Gst-Ap12 bacterial lysate was used. A similar pattern of complex formation was also observed by Helin et al. (26). Later experiments showed that this discrepancy was due to the amount of protein used in the binding assay, since an increased concentration of Ap12 protein in the binding reaction resulted in formation of two DNA-protein complexes (data not shown; Fig. 7C, lanes 11 and 12). The dose dependency of the second complex could be due to the ability of the fusion products to form a dimer or multimer or simply to the fact that the DNA probe used for the mobility shift assay contains two E2F recognition sites.

To determine whether RB is able to interact with the specific Ap12-DNA complex, purified p56-RB protein was included in the DNA mobility shift assay. The experiments were performed in two ways: either Gst-SH5 was mixed with RB and then added to the E2F probe (Fig. 7B, lane 3), or the fusion protein was first bound to the E2F probe and then RB was added (Fig. 7B, lane 6). In either case, the Ap12-DNA complex was shifted to a more slowly migrating position, indicating that RB has the ability to interact with the specific Ap12-DNA complex. These results strongly suggest that the Ap12 protein has DNA-binding, as well as RB-binding, properties similar to those of E2F. Similar conclusions were reached by other two groups (26, 35).

To map the region responsible for DNA binding, various deletion or truncation mutants of Ap12-Gst fusions (Fig. 7D) were constructed, expressed in *E. coli*, and assayed for DNA-binding activity. As shown in Fig. 7C, SH5 and XH9 bound to the E2F recognition sequence whereas the C-terminal region of Ap12 (P3) did not. This indicated that the DNA-binding domain of Ap12 was located between amino acids 162 and 362. Two additional N-terminal deletions, BB8 and SA1, did not show the DNA-binding activity, suggesting that the N-terminal limits of the sequences required for DNA-binding reside at amino acid 162. To map the DNA-binding region more precisely, fusion proteins of BH7 and HH1 were assayed for DNA-binding activity. Interestingly, BH7 (amino acids 162 to 263) retained significant DNA-binding activity while deletion of amino acids 162 to 190 and 244 to 263 in HH1 (Fig. 7C and D) abolished the ability to bind DNA. These results demonstrate that the region between amino acids 162 and 263, including the putative leucine repeats (amino acids 192 to 213) and flanking basic amino acids, confers the ability of the Ap12 fusion protein to bind the E2F recognition sequence. These data agree with the finding of Helin et al. (26) that the DNA-binding domain maps to the region between amino acids 89 and 191 (or amino acids 128 and 230, by our numbering method). Taken together, the results suggest that the sequences required for the DNA binding reside within amino acids 162 to 230 (our numbering), which contain the putative leucine repeats.

**Expression of Ap12 in CV1 cells transactivates a promoter with E2F recognition sequences.** To determine whether Ap12 can activate transcription from an E2F-binding site-dependent promoter, plasmids CMVAp12Stu and CMVAp12RH (Fig. 8A) were constructed to express Ap12 in mammalian cells under the control of a cytomegalovirus immediate-early promoter. In addition, CMVAp12ΔDS, which was derived from CMVAp12Stu and contains an internal deletion (Δ343-568), was constructed to serve as an additional control. Three reporter plasmids, pE2FA$_{10}$CAT, pE2FmA$_{10}$CAT, and pA$_{10}$CAT, containing, respectively, a wild-type, a mutated, and no E2F-binding site upstream of the CAT reporter gene (68), were used for this assay. Figure 8B shows that expression of either CMVAp12Stu or CMVAp12RH significantly enhanced CAT activity when cotransfected with pE2FA$_{10}$CAT but not when cotransfected with pE2FmA$_{10}$CAT or pA$_{10}$CAT, whereas transfection of CMVAp12ΔDS or cytomegalovirus alone had no effect on CAT activity. These data demonstrate that Ap12 encodes a functional transcription factor that activates promoters with E2F recognition sequences.

## DISCUSSION

**Identification of RbAps.** A simplified model for RB function is that relatively few target molecules which play central roles in cellular function are regulated by the RB protein.

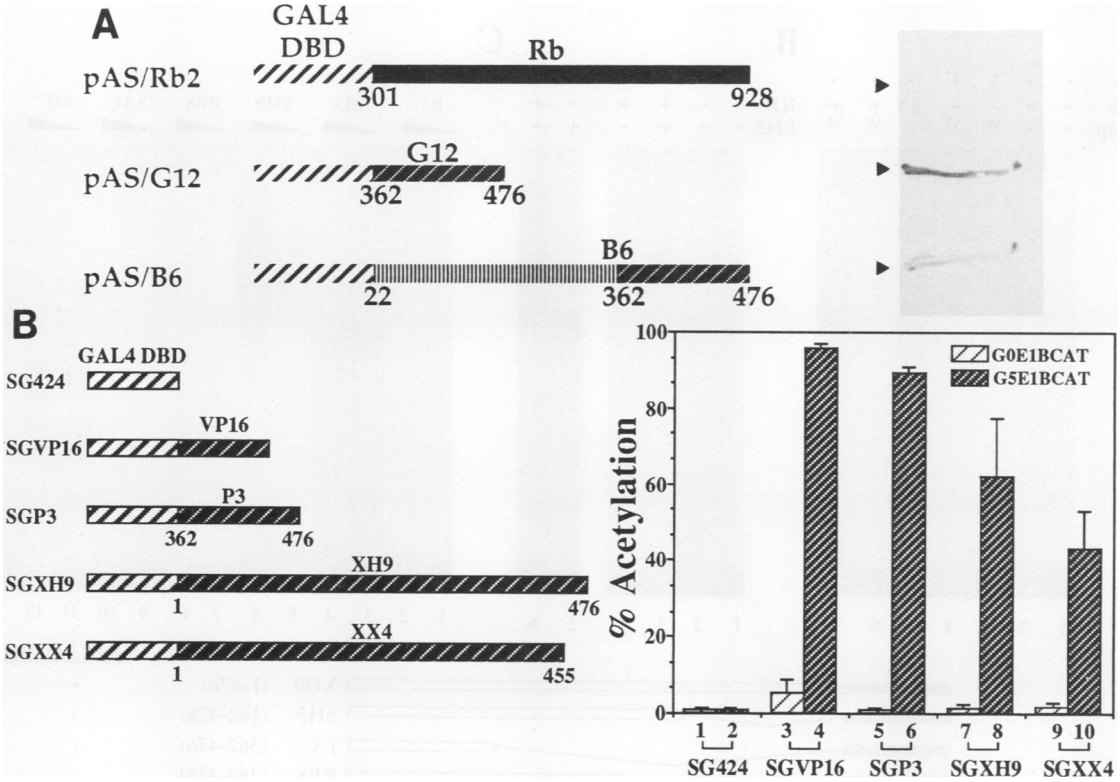



FIG. 6. The C terminus of Ap12 serves as an activation domain when fused to the GAL4 DNA-binding domain as assayed in both yeast and mammalian cells. (A) Fusion proteins of the GAL4 DNA-binding domain (DBD) (amino acids 1 to 147) with either Rb2 (RB amino acids 301 to 928), Ap12-G12 (amino acids 362 to 476), or Ap12-B6 (amino acids 22 to 476) were constructed as shown in the left half of panel A and were expressed in *S. cerevisiae* by using the pAS1 vector as previously described (15). A single colony of each transformant was streaked onto plates and further analyzed for β-galactosidase activity by using a colony lift assay as shown in the right panel. The G12 and B6 yeast transformants were blue, indicating that the expressed fusion proteins act as a GAL4 transactivation factor and the fused polypeptide may contain transactivation sequences. Expression of the RB fusion protein in negative control plasmid pAS/Rb2 was detected by Western blotting analysis. The arrowheads indicate the positions of the proteins to the left. Similar results obtained with mammalian cells are shown in panel B. Three different fragments of Ap12 excised from Ap12-Gst fusions (P3, XH9, and XX4) were cloned into the pSG424 expression vector, which contains a GAL4 DNA-binding domain (left panel). These GAL4 chimeric expression plasmids were cotransfected with E1BCAT reporter genes harboring no or five GAL4-binding sites, and CAT activities (right panel) were quantitated with a Betascope 603 Blot Analyzer (Betagen, Boston, Mass.). Plasmids pSG424 and pSGVP16 were included as controls.

Inactivation of RB by any one of three means, phosphorylation (11, 13), mutation (28, 58), or oncoprotein binding (12, 22, 67), could potentially uncouple RB connections and lead to deregulated growth. Until this report, only a limited number of molecules were known to be capable of interacting with RB, such as two proteins of unknown function, p1 and p2, the *myc* protein, and 8 to 10 other, unidentified proteins. To elucidate genetically and biochemically how RB functions in a cell, it is essential to identify as many of the genes that encode potential partners of RB as possible. The approach described in this report is to use an RB sandwich to screen λgt11 cDNA expression libraries. The advantages of this one-step RB sandwich procedure are its simplicity and directness and that the clone isolated should encode a fusion protein that would directly interact with RB in the absence of potential bridging proteins. A likely disadvantage of this approach is that some associated proteins would be missed because the interaction site might be masked by the RB antibody. Since we were most interested in isolating clones having a binding specificity similar to that of T antigen, screening was performed with SV40 large-T antigen as a positive control. A λgt11 phage expressing T antigen was constructed for this purpose, and the association between RB and T can be readily detected by this method, indicating that the sandwich method of screening would work. The cloning method we described was based on use of purified RB protein as a tag to screen cDNA expression libraries, similar to the methods described by two other groups (26, 35). As a consequence, all three laboratories have cloned what appears to be the same gene that encodes a protein with properties similar to those of transcription factor E2F.

In addition to Ap12, eight other distinct cDNA clones were also isolated. Immunostaining and cell fractionation showed that all of the gene products examined were located in the nucleus (Table 1). This is an important criterion for any protein that interacts with RB in a biologically significant manner, since the interaction probably would occur in the nucleus (44). Interestingly, one of the clones appeared to be the gene for E7 of HPV from the HeLa cell cDNA library. The fact that E7 is one of the known viral proteins bound by RB (16) strengthens greatly the accuracy and utility of the RB sandwich screening method. Although we have not established the biological relevance of the RB associations with the other two known proteins (upstream binding factor and lamin C), the interaction with proteins that have such

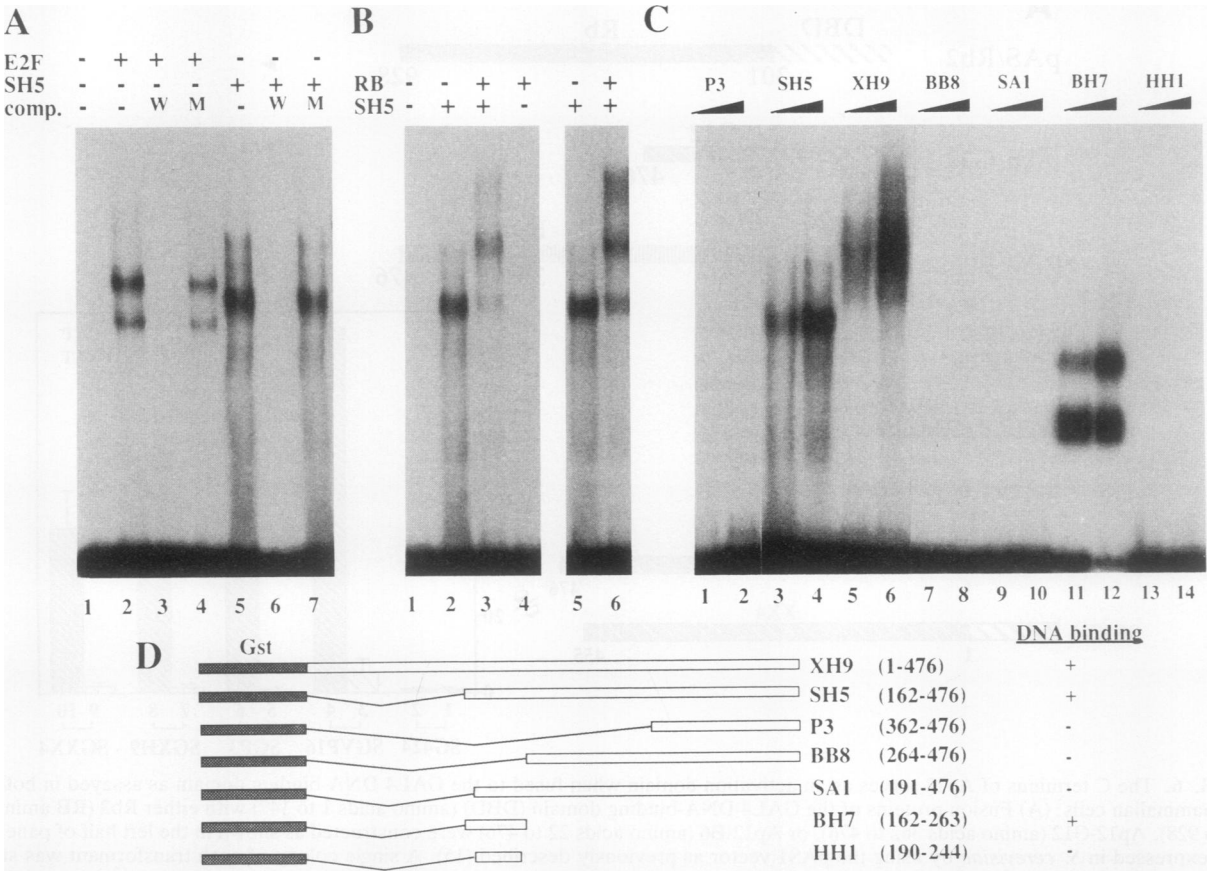



FIG. 7. Ap12 binds specifically to the E2F recognition sequence. Fusion proteins of Gst-Ap12 (P3, SH5, XH9, BB8, SA1, BH7, and HH1) were expressed in *E. coli*, and the bacterial lysates were used for a DNA mobility shift assay. The probe was a DNA fragment containing two E2F recognition sites which was $^{32}P$ end labeled by a Klenow fill-in reaction. (A) Gst-Ap12SH5 binds to the E2F-specific sequence. As a positive control, a partially purified E2F protein from HeLa cells was also used (65). Approximately 1 μg of a DNA fragment containing either the wild-type E2F sites or mutated E2F sites were used as a competitor. Lanes: 1, probe alone; 2, E2F plus probe; 3, E2F plus probe plus wild-type competitor; 4, E2F plus probe plus mutant competitor; 5, SH5 plus probe; 6, SH5 plus probe plus wild-type competitor; 7, SH5 plus probe plus mutant competitor. (B) RB interacts with the Ap12-E2F DNA complex. Lanes: 1, probe alone; 2, SH5 plus probe; 3, SH5 plus p56-RB (0.25 μg), with preincubation for 15 min followed by probe addition; 4, p56-RB plus probe; 5, SH5 plus probe; 6, SH5 plus probe incubated for 15 min, then p56-RB added. (C) The DNA-binding domain of Ap12 is located at a region containing a leucine repeat motif. Two different amounts of each fusion protein were used. Lanes: 1, P3, 200 ng; 2, P3, 400 ng; 3, SH5, 20 ng; 4, SH5, 40 ng; 5, XH9, 20 ng; 6, XH9, 40 ng; 7, BB8, 100 ng; 8, BB8 200 ng; 9, SA1, 100 ng; 10, SA1, 200 ng; 11, BH7, 20 ng; 12, BH7, 40 ng; 13, HH1, 100 ng; 14, HH1, 200 ng. Quantities of the fusion proteins were estimated on the basis of Coomassie blue staining. (D) Diagrams of constructs used to map the DNA-binding region of Ap12. The leucine repeat motif is located at amino acids 192 to 213. Fusion proteins XH9, SH5, and BH7, containing the repeat and its N-terminal flanked amino acids, were able to bind the DNA (+), while the rest of the fusion proteins were not (−).

diverse functions reveals an unexpected degree of complexity with regard to RB function.

**Transcription factors as targets of regulation by the RB protein.** In our collection of nine newly cloned RbAps, one is a known eukaryotic upstream binding factor which recognizes and binds to the rRNA promoter and activates transcription mediated by RNA polymerase I through cooperative interactions with SL1 (33), and another, Ap12, has properties consistent with those proposed for the E2F transcription factor. The C-terminal region of Ap12, which contains a transactivation domain, is very acidic, a hallmark of the transactivation domains of several known transcription factors, such as GAL4 and VP16 (49, 54). The DNA-binding domain appears to be located in the middle region of the protein, which features a putative leucine zipper motif flanked by stretches of basic amino acids (Fig. 9). The precise basic region for DNA binding remains to be elucidated by a more detailed analysis. Most significantly, we have shown that the cloned gene product has the E2F site-dependent transcriptional activation activity. Similar results were also reported by Helin et al. (26), although they used a plasmid that expresses a protein from which 87 amino acids of the N terminus were truncated. All three laboratories have shown that the cloned gene product has the properties of transcription factor E2F. However, whether this cloned gene is identical to E2F (9, 68) remains unclear. Numerous trials using antibodies raised against the Ap12 gene product failed either to supershift an E2F-DNA complex or to preabsorb the E2F-binding activity. Similar conclusions were also reached by Helin et al. (26), suggesting the possible existence of multiple members in the E2F family.

If the cellular function of RB is to restrict entry of cells into $G_1$ (20), the genes important for $G_1$ progression and entrance into the S phase are most likely to be regulated

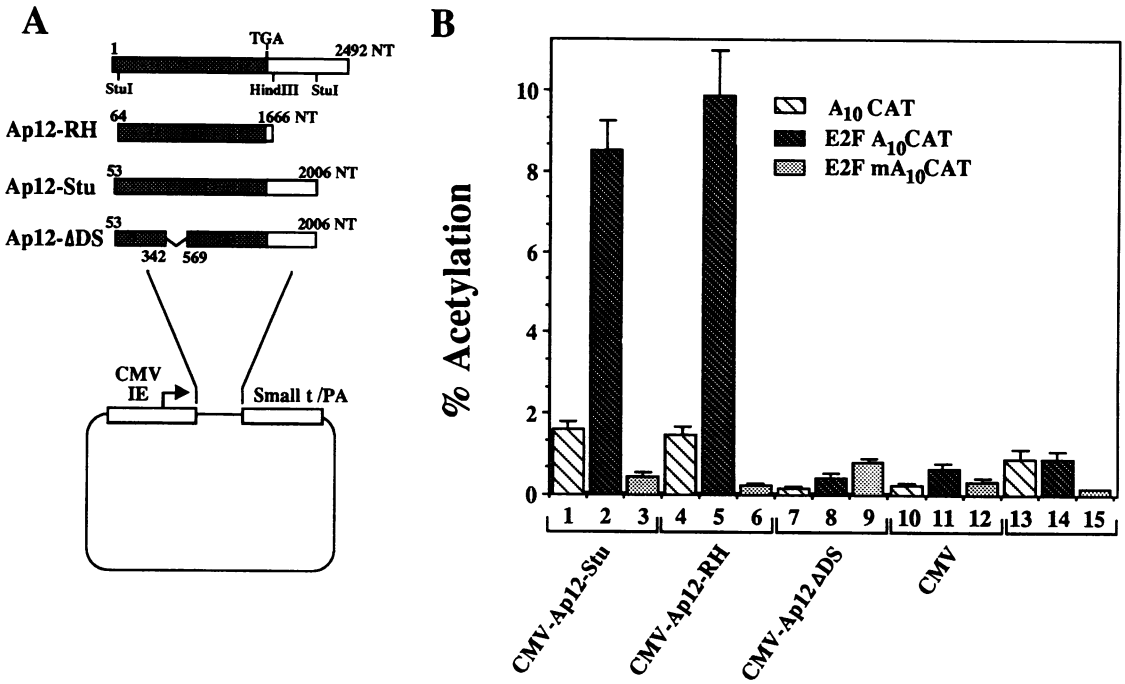



FIG. 8. Ap12 transactivates a promoter with an E2F recognition site. (A) Diagrams of the constructs of Ap12 expression vectors. The promoter of the expression vector is cytomegalovirus immediate-early gene promoter (CMV-IE), and the 3′ polyadenylation site is derived from small t/PA. Two different Ap12 cDNAs (*StuI* fragment from A6 and *Eco*RI-*Hin*dIII fragment from B6) were inserted into this vector. CMV-Ap12ΔDS is derived from CMV-Ap12Stu and contains a truncation deletion of the *Dra*III-*Sal*I fragment. NT, nucleotides. (B) Transcriptional activation of a promoter with E2F recognition sequences. Five micrograms each of reporter plasmids $pA_{10}CAT$, $pE2FmA_{10}CAT$, and $pE2FA_{10}CAT$ was cotransfected with 5 μg of CMV-Ap12Stu, CMV-Ap12RH, and CMV-Ap12ΔDS, respectively, or the vector alone was transfected into monkey kidney CV1 cells. The cells were harvested after 48 h, and CAT activities were measured as previously described (21). Transfection with the reporter plasmids alone (−) served as controls. Transfection efficiency was standardized with β-galactosidase activity by cotransfection with a plasmid expressing the *lacZ* gene.

directly or indirectly by RB. Transcription factor E2F is known to associate with RB in a cell cycle-dependent manner (8, 50, 59), with a tight association prevalent in the $G_0$-$G_1$ stage but not in the S or M phase. It is reasonable to propose that RB sequesters E2F in the $G_0$-$G_1$ stage in an inactive conformation. Its release from the RB complex somehow allows it to assume an active conformation that is capable of influencing its target genes through interactions with E2F DNA-binding sites and the general transcriptional machinery. It has been reported that there are several genes, including *myc*, the dihydrofolate reductase gene, and *myb*, that may be subject to E2F transcriptional control (27, 51). Our preliminary results indicate that the level of Ap12 mRNA in serum-deprived cells was increased between 2 to 6 h after serum stimulation (56a), which coincides with the pattern of expression of delayed-early growth response genes (40). It is likely that other members of the E2F family could play a role in controlling the expression of those immediate-early genes. An important challenge is to determine the identity of the Ap12-E2F target genes and to ascertain their roles in the control of the cell cycle. While the level of Ap12-E2F mRNA appears to be cell cycle regulated, the activity of Ap12-E2F could be subject to multilevel regulation, including association with RB, phosphorylation, and protein degradation. It is plausible that the precise amounts of Ap12-E2F and RB expressed during the $G_0$-$G_1$-to-S transition of the cell cycle and their phosphorylation status are critical for regulation of cell cycle progression.

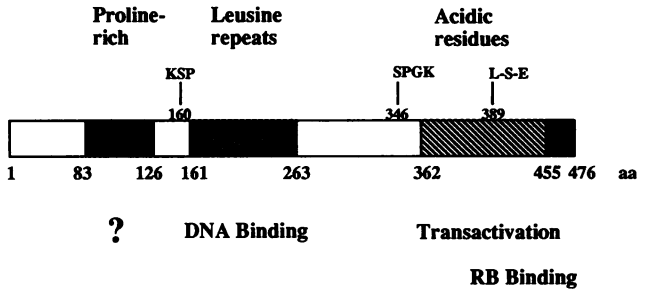



FIG. 9. Schematic diagram of Ap12 showing its structural-functional relationship. The first amino acid was designated position 1. The putative functional domains of the protein are shaded. The proline-rich region with an unknown function is located at amino acids (aa) 83 to 126. The region from amino acids 161 to 263 is a DNA-binding domain containing leucine repeats. The acidic C-terminal region has the transactivation activity. The fragment from amino acids 455 to 476 is required for RB binding. Two putative kinase recognition sites are also indicated.

## ACKNOWLEDGMENTS

We thank Amy Yee for providing all of the valuable E2F reagents and for constructive advice during this work. Thanks also to Barbara Christy for comments and to all of the members of the Lee laboratory for contributions to this project.

This work is supported by grants from the NIH (EY 05758) and the Council for Tobacco Research to W.H.L.

## REFERENCES


1. **Bagchi, S., R. Weinmann, and P. Raychaudhuri.** 1991. The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. Cell **65**:1063–1072.
2. **Bandara, L. R., J. P. Adamczewski, T. Hunt, and N. B. LaThangue.** 1991. Cyclin A and the retinoblastoma gene product complex with a common transcription factor. Nature (London) **352**:249–251.
3. **Bookstein, R., and W.-H. Lee.** 1991. Molecular genetics of the retinoblastoma suppressor gene. Crit. Rev. Oncogenesis **2**:211–227.
4. **Bookstein, R., P. Rio, S. Madreperla, F. Hong, C. Allred, W. E. Grizzle, and W.-H. Lee.** 1990. Promoter deletion and loss of retinoblastoma gene expression in human prostate carcinoma. Proc. Natl. Acad. Sci. USA **87**:7762–7766.
5. **Bookstein, R., J.-Y. Shew, P.-L. Chen, P. Scully, and W.-H. Lee.** 1990. Suppression of tumorigenicity of human prostate carcinoma cells by replacing a mutated *RB* gene. Science **247**:712–715.
6. **Breeden, L., and K. Nasmyth.** 1985. Regulation of the yeast HO gene. Cold Spring Harbor Symp. Quant. Biol. **50**:643–650.
7. **Buchkovich, K., L. A. Duffy, and E. Harlow.** 1989. The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. Cell **58**:1097–1105.
8. **Cao, L., B. Faha, M. Dembski, L.-H. Tsai, E. Harlow, and N. Dyson.** 1992. Independent binding of the retinoblastoma protein and p107 to the transfection factor E2F. Nature (London) **355**:176–179.
9. **Chellappan, S. P., S. Hiebert, M. Mudryj, J. M. Horowitz, and J. R. Nevins.** 1991. The E2F transcription factor is a cellular target for the RB protein. Cell **65**:1053–1061.

9a. **Chen, P.-L.** Unpublished data.

10. **Chen, P.-L., Y. Chen, B. Shan, R. Bookstein, and W.-H. Lee.** 1992. Stability of RB expression determines the tumorigenicity of reconstituted retinoblastoma cells. Cell Growth Differ. **3**:119–125.
11. **Chen, P.-L., P. Scully, J.-Y. Shew, J. Y.-J. Wang, and W.-H. Lee.** 1989. Phosphorylation of the retinoblastoma gene product is modulated during the cell cycle and cellular differentiation. Cell **58**:1193–1198.
12. **DeCaprio, J. A., J. W. Ludlow, J. Figge, J.-Y. Shew, C.-M. Huang, W.-H. Lee, E. Marsillo, E. Paucha, and D. M. Livingston.** 1988. SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. Cell **54**:275–283.
13. **DeCaprio, J. A., J. W. Ludlow, D. Lynch, Y. Furukawa, J. Griffin, H. Pawnica-Worms, C.-M. Huang, and D. M. Livingston.** 1989. The product of the retinoblastoma susceptibility gene has properties of a cell cycle regulatory element. Cell **58**:1085–1095.
14. **Defeo-Jones, D., P. S. Huang, R. E. Jones, K. M. Haskell, G. A. Vuocolo, M. G. Hanobik, H. E. Huber, and A. Oliff.** 1991. Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product. Nature (London) **352**:251–254.
15. **Devoto, S. H., M. Mudryj, J. Pines, T. Hunter, and J. R. Nevins.** 1992. A cyclin A protein kinase complex possesses sequence-specific DNA binding activity: $p33^{cdk2}$ is a component of the E2F-cyclin A complex. Cell **68**:167–176.

15a. **Durfee, T., et al.** Unpublished data.

16. **Dyson, N., P. M. Howley, K. Munger, and E. Harlow.** 1989. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science **243**:934–937.
17. **Fields, S., and S. K. Jang.** 1990. Presence of a potent transcription activating sequence in the p53 protein. Science **249**:1046–1051.
18. **Fisher, D. Z., N. Chaudhary, and G. Blobel.** 1986. cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. Proc. Natl. Acad. Sci. USA **83**:6450–6454.
19. **Friend, S. H., R. Bernards, S. Rogelj, R. A. Weinberg, J. M. Rapaport, D. M. Albert, and T. P. Dryja.** 1986. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature (London) **323**:643–646.
20. **Fung, Y. K. T., A. L. Murphree, A. T'Ang, J. Qian, S. H. Hinrichs, and W. F. Benedict.** 1987. Structural evidence for the authenticity of the human retinoblastoma gene. Science **236**:1657–1661.
21. **Goodrich, D. W., Y. Chen, P. Scully, and W.-H. Lee.** 1992. Expression of the retinoblastoma gene product in bladder carcinoma cells associates with a low frequency of tumor formation. Cancer Res. **52**:1968–1973.
22. **Goodrich, D. W., N. P. Wang, Y.-W. Qian, E. Y.-H. P. Lee, and W.-H. Lee.** 1991. The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle. Cell **67**:293–302.
23. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
24. **Gullemont, F., A. Billault, and C. Auffray.** 1989. Physical linkage of guanine nucleotide-binding protein-related gene to the chicken major histocompatibility complex. Proc. Natl. Acad. Sci. USA **86**:4594–4598.
25. **Harbour, J. W., S.-H. Lai, J. Whang-Peng, A. F. Gazdar, J. D. Minna, and F. J. Kaye.** 1988. Abnormalities in structure and expression of the human retinoblastoma gene in SCLC. Science **241**:353–357.
26. **Helin, K., J. A. Lees, M. Vidal, N. Dyson, E. Harlow, and A. Fattaey.** 1992. A cDNA encoding a pRB-binding protein with properties of the transcription factor E2F. Cell **70**:337–350.
27. **Hiebert, S. W., M. Lipp, and J. R. Nevins.** 1989. E1A-dependent trans-activation of the human MYC promoter is mediated by the E2F factor. Proc. Natl. Acad. Sci. USA **86**:3594–3598.
28. **Horowitz, J. M., D. W. Yandell, S. H. Park, S. Canning, P. Whyte, K. Buchkovich, E. Harlow, R. Weinberg, and T. Dryja.** 1989. Point mutational inactivation of the retinoblastoma antioncogene. Science **243**:937–940.
29. **Hu, Q., N. Dyson, and E. Harlow.** 1990. The regions of the retinoblastoma protein needed for binding to adenovirus E1A or SV40 T antigen are common sites for mutations. EMBO J. **9**:1147–1155.
30. **Huang, H.-J. S., J.-K. Yee, J.-Y. Shew, P.-L. Chen, R. Bookstein, T. Friedmann, E. Y.-H. P. Lee, and W.-H. Lee.** 1988. Suppression of the neoplastic phenotype by replacement of the retinoblastoma gene product in human cancer cells. Science **242**:1563–1566.
31. **Huang, S., W.-H. Lee, and E. Y.-H. P. Lee.** 1991. Identification of a cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product. Nature (London) **350**:160–162.
32. **Huang, S., E. Shin, K.-A. Sheppard, L. Chokroverty, B. Shan, Y.-W. Qian, E. Y.-H. P. Lee, and A. S. Yee.** 1992. The retinoblastoma protein region required for interaction with the E2F transcription factor includes the T/E1A binding and C-terminal sequences. DNA Cell Biol. **11**:539–548.
33. **Huang, S., N.-P. Wang, B. Y. Tseng, W.-H. Lee, and E. Y.-H. P. Lee.** 1990. Two distinct and frequently mutated regions of retinoblastoma protein are required for binding to SV40 T antigen. EMBO J. **9**:1815–1822.
34. **Jantzen, H.-M., A. Admon, S. P. Bell, and R. Tjian.** 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature (London) **344**:830–836.
35. **Kaelin, W. G. J., W. Krek, W. R. Sellers, J. A. DeCaprio, F. Ajchenbaum, C. S. Fuchs, T. Chittenden, Y. Li, P. J. Farnham, M. A. Blanar, D. M. Livingston, and E. K. Flemington.** 1992. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. Cell **70**:351–364.
36. **Kaelin, W. G. J., D. C. Pallas, J. A. DeCaprio, F. J. Kaye, and D. M. Livingston.** 1991. Identification of cellular proteins that can interact specifically with the T/E1A-binding region of the retinoblastoma gene product. Cell **64**:521–532.
37. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science **231**:699–704.
38. **Keyomarsi, K., L. Sandoval, V. Band, and A. B. Pardee.** 1991. Synchronization of tumor and normal cells from G1 to multiple

cell cycles by lovastatin. Cancer Res. **51**:3602–3609.
39. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**:1759–1764.
40. **Lau, L. F., and D. Nathans.** 1991. Genes induced by serum growth factors, p. 257–293. *In* P. Cohen and J. G. Foulkes (ed.), The hormonal control regulation of gene transcription. Elsevier Science Publishers, Amsterdam.
41. **Lee, E. Y.-H. P., H. To, J.-Y. Shew, R. Bookstein, P. Scully, and W.-H. Lee.** 1988. Inactivation of the retinoblastoma susceptibility gene in human breast cancers. Science **241**:218–221.
41a. **Lee, W.-H., et al.** Unpublished data.
42. **Lee, W.-H., R. Bookstein, F. Hong, L.-J. Young, J.-Y. Shew, and E. Y.-H. P. Lee.** 1987. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. Science **235**:1394–1399.
43. **Lee, W.-H., R. E. Hollingsworth, Y.-W. Qian, P.-L. Chen, F. Hong, and E. Y.-H. P. Lee.** 1991. The RB protein as a cellular "corral" for growth-promoting proteins. Cold Spring Harbor Symp. Quanti. Biol. **61**:211–217.
44. **Lee, W.-H., J.-Y. Shew, F. Hong, T. Sery, L. A. Donoso, L. J. Young, R. Bookstein, and E. Y.-H. P. Lee.** 1987. The retinoblastoma susceptibility gene product is a nuclear phosphoprotein associated with DNA binding activity. Nature (London) **329**: 642–645.
45. **Lees, J. A., K. J. Buchkovich, D. R. Marshak, C. W. Anderson, and E. Harlow.** 1991. The retinoblastoma protein is phosphorylated on multiple sites by human cdc2. EMBO J. **10**:4279–4290.
46. **Lillie, J. W., and M. R. Green.** 1989. Transcription activation by the adenovirus E1a protein. Nature (London) **338**:39–44.
47. **Lin, B. T., S. Gruenwald, A. O. Morla, W.-H. Lee, and J. Y. Wang.** 1991. Retinoblastoma cancer suppressor gene product is a substrate of the cell cycle regulator cdc2 kinase. EMBO J. **10**:857–864.
48. **McKeon, F., M. Kirschner, and D. Caput.** 1986. Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. Nature (London) **319**: 463–468.
49. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
50. **Mudryj, M., S. H. Devoto, S. W. Hiebert, T. Hunter, J. Pines, and J. R. Nevins.** 1991. Cell cycle regulation of the E2F transcription factor involves an interaction with cyclin A. Cell **28**:1243–1253.
51. **Mudryj, M., S. W. Hiebert, and J. R. Nevins.** 1990. A role for the adenovirus inducible E2F transcription factor in a proliferation dependent signal transduction pathway. EMBO J. **9**:2179–2184.
52. **Phelps, W. C., K. Munger, C. L. Yee, J. A. Barnes, and P. M. Howley.** 1992. Structure-function analysis of the human papillomavirus type 16 E7 oncoprotein. J. Virol. **66**:2418–2427.
53. **Rustgi, A. K., N. Dyson, and R. Bernards.** 1991. Amino-terminal domains of c-myc and N-myc proteins mediate binding to the retinoblastoma gene product. Nature (London) **352**:541–544.
54. **Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne.** 1988. GAL4-VP16 is an unusually potent transcriptional activator. Nature (London) **335**:563–564.
55. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
56. **Schiestl, R. H., and R. D. Gietz.** 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. **16**:339–346.
56a. **Shan, B.** Unpublished data.
57. **Shenoy, S., J.-K. Choi, S. Bagrodia, T. D. Copeland, J. L. Maller, and D. Shalloway.** 1989. Purified maturation promoting factor phosphorylates pp60c-src at the sites phosphorylated during fibroblast mitosis. Cell **57**:763–774.
58. **Shew, J.-Y., B. Lin, P.-L. Chen, B. Y. Tseng, T. L. Yang-Feng, and W.-H. Lee.** 1990. C-terminal truncation of the RB protein leads to functional inactivation. Proc. Natl. Acad. Sci. USA **87**:6–10.
59. **Shirodkar, S., M. Ewen, J. A. Decaprio, J. Morgan, D. M. Livingston, and T. Chittenden.** 1992. The transcription factor E2F interacts with the retinoblastoma product and a p107-cyclin A complex in a cell cycle-regulated manner. Cell **68**:157–166.
60. **Smith, D. B., and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in E. coli as fusions with glutathione S-transferase. Gene **67**:31–40.
61. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60–89.
62. **Sumegi, J., E. Uzvolgyi, and G. Klein.** 1990. Expression of the RB gene under the control of MuLV-LTR suppresses tumorigenicity of WERI-Rb-27 retinoblastoma cells in immunodefective mice. Cell Growth Differ. **1**:247–250.
63. **Takahashi, R., T. Hashimoto, X. Hong-Ji, S.-X. Hu, T. Matsui, T. Miki, H. Bigo-Marshall, S. A. Aaronson, and W. F. Benedict.** 1991. The retinoblastoma gene functions as a growth and tumor suppressor in human bladder carcinoma cells. Proc. Natl. Acad. Sci. USA **88**:5257–5261.
64. **Toguchida, J., K. Ishizaki, M. S. Sasaki, M. Ikenaga, M. Sugimoto, Y. Kotoura, and T. Yamamuro.** 1988. Chromosomal reorganization for the expression of recessive mutation of retinoblastoma susceptibility gene in the development of osteosarcoma. Cancer Res. **48**:3939–3943.
65. **Vinson, C. R., P. B. Sigler, and S. L. McKnight.** 1989. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science **246**:911–916.
66. **Wang, N.-P., P.-L. Chen, S. Huang, L. A. Donoso, W.-H. Lee, and E. Y.-H. P. Lee.** 1990. DNA-binding activity of retinoblastoma protein is intrinsic to its carboxyl-terminal region. Cell Growth Differ. **1**:233–239.
67. **Whyte, P., K. J. Buchkovich, J. M. Horowitz, S. H. Friend, M. Raybuck, R. A. Weinberg, and E. Harlow.** 1988. Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature (London) **334**:124–129.
68. **Yee, A. S., P. Raychaudhuri, L. Jakoi, and J. R. Nevins.** 1989. The adenovirus-inducible factor E2F stimulates transcription after specific DNA binding. Mol. Cell. Biol. **9**:578–585.
69. **Yokota, J., T. Akiyama, Y.-K. T. Fung, W. F. Benedict, Y. Namba, M. Hanaoka, M. Wada, T. Terasaki, Y. Shimosato, T. Sugimura, and M. Terada.** 1988. Altered expression of the retinoblastoma (RB) gene in small-cell carcinoma of the lung. Oncogene **3**:471–475.