Identification of Discrete Structural Domains in the Retinoblastoma Protein

AMINO-TERMINAL DOMAIN IS REQUIRED FOR ITS OLIGOMERIZATION*

(Received for publication, August 19, 1993)

Carmel E. Hensey, Frank Hong, Tim Durfee, Yue-Wei Qian, Eva Y.-H.P. Lee, and Wen-Hwa Lee‡
From the Center for Molecular Medicine/Institute of Biotechnology, University of Texas Health Science Center,
San Antonio, Texas 78245

To characterize the protein product of the retinoblastoma tumor suppressor gene biochemically, a recombinant human protein was produced in an Escherichia coli expression system. The full-length protein, p110\textsuperscript{N}, and an amino-terminal truncated form, p56\textsuperscript{N}, were expressed and purified to near homogeneity by conventional chromatographic procedures. To probe the structural organization of the retinoblastoma protein the purified proteins were subjected to partial proteolysis by trypsin, chymotrypsin, and subtilisin. Four discrete structural domains were revealed in p110\textsuperscript{N} by this method. Two of these structural domains, found in both p56\textsuperscript{N} and p110\textsuperscript{N}, were mapped to the carboxy-terminal half of the protein and corresponded to the SV40 large T binding domains defined previously by genetic methods. In addition two distinct domains in the amino-terminal half of the protein were also defined. A potential role for these newly defined amino-terminal domains was uncovered upon analysis of the purified proteins by non-denaturing polyacrylamide gel electrophoresis. p110\textsuperscript{N} revealed multiple bands by this method, suggesting the formation of oligomeric structures by the protein, while this property was not observed for p56\textsuperscript{N}. Electron microscopy of p110\textsuperscript{N} revealed linearly extended, macromolecular structures, further supporting the formation of homologous higher order structures by the full-length retinoblastoma protein. Analysis of the interactions between retinoblastoma protein molecules using the yeast two-hybrid system confirmed that the retinoblastoma protein contains multiple structural domains with the amino-terminal domains being required for oligomerization of the full-length protein.

*To whom correspondence should be addressed. Tel.: 210-677-6350; Fax: 210-677-6377.
‡The abbreviations used are: RB, retinoblastoma protein; PMSF, phenylmethysulfonyl fluoride; PAGE, polyacrylamide gel electrophoresis; TPC, l-1-tosylamido-2-phenylethyl chloromethyl ketone.

investigation in an effort to unravel its role in the regulation of cell growth (Goodrich and Lee, 1993b; Hollingsworth, et al., 1993). The protein was identified upon cloning of the gene whose defect predisposes to the pediatric cancer, retinoblastoma (Friend et al., 1986; Lee et al., 1987b), and was subsequently found to be mutated in many cancers including osteosarcoma and small cell lung carcinoma (Harbor et al., 1988; Horowitz et al., 1990; Shew et al., 1989). Reintroduction of the wild-type RB gene into these tumor cells, by retrovirus-mediated gene transfer, consistently suppressed their tumorigenic ability, suggesting that RB is a general tumor suppressor (Huang et al., 1988; Lee and Lee, 1991).

RB is a nuclear phosphoprotein that exhibits cell cycle-dependent changes in its phosphorylation state (Buchkovitch et al., 1989; Chen et al., 1989; DeCaprio et al., 1989), this being the first indication that RB may play a role in the cell division cycle. Microinjection of the full-length or carboxy-terminal half of the protein into cells inhibits G1 phase progression, suggesting that RB may function during early G1 (Goodrich et al., 1991). The RB protein has been shown to form specific complexes with the transforming proteins of several DNA tumor viruses, including SV40 large T, adenovirus E1A, and human papilloma virus E7 (DeCaprio et al., 1988; Dyson et al., 1989; Whyte et al., 1988). Deletion analysis indicates the existence of two noncontiguous regions, from amino acids 394 to 571 (domain A) and from 649 to 773 (domain B), both of which are necessary for the binding of these transforming proteins (Huang et al., 1990; Kaelin et al., 1990). Recently, this region of the RB protein was found to bind many cellular proteins including the transcription factor, E2F-1 (Helin et al., 1992; Kaelin et al., 1992; Shan et al., 1992), p48 (Qian et al., 1993), RBP-1, RBP-2 (Defeo-Jones et al., 1991), RhAP46 (Huang et al., 1991), and a protein phosphatase type 1 catalytic subunit (Durfee et al., 1993). This is also the region that suffers the most mutations found in human tumors (Bookstein and Lee, 1991). The carboxy-terminal one-third of the RB protein contains both nuclear localization sequences (Shew et al., 1990) and a DNA binding activity, which to date has not been shown to be specific for any particular DNA sequence (Lee et al., 1987a; Wang et al., 1990b).

Availability of pure proteins in sufficient quantities is an important prerequisite for detailed structural and functional analyses of proteins. Here we report the production and purification of both p110\textsuperscript{N} and p56\textsuperscript{N} from an Escherichia coil expression system in quantities sufficient for biochemical and structural analysis. It is widely recognized that globular proteins are composed of discrete globular regions, generally referred to as domains (Rose, 1979; Wetlaufer, 1973; Wilson, 1992). Limited proteolysis has frequently been found to result in cleavage between structural domains, as the relatively
unstructured segments of polypeptide chain connecting domains are usually quite susceptible to proteolytic digestion (Fontana et al., 1986). In the present study we have analyzed the products of limited proteolysis of purified p110RB and p56aa, using the proteolytic enzymes, trypsin, chymotrypsin, and subtilisin. Four structural domains were identified in the protein by this method, the location of which was determined by microsequence analysis.

The structure and function of the amino-terminal half of the RB protein are largely uncharacterized. No naturally occurring domain has been found to contain mutations in this region. However, two of the structural domains identified mapped to this portion of the protein, and a cellular protein that specifically binds to this region has been identified.2 In addition, while characterizing the full-length RB protein a novel biochemical property was observed. Specifically, it concerns the ability of the purified RB protein to form oligomers in vitro, a property that requires the presence of its amino-terminal region. This additional level of structural organization between RB molecules suggests complex structure/function relationships for this protein and a possible functional significance for the amino-terminal domain. These results provided further insights into the structure of the RB protein thus helping to establish structure/function relationships within the protein.

MATERIALS AND METHODS

Cloning and Expression of Retinoblastoma Proteins—Retinoblastoma proteins were expressed in E. coli using the pET8c construct previously described (Huang et al., 1991). The entire coding sequence of Rb cDNA (2.8 kilobase pairs) was placed under the control of the T7 polymerase promoter using the pET8c vector (Rosenberg et al., 1987). Two species of RB were expressed from this construct, a full-length p109RB protein and p56aa, an amino-terminal truncated form of the protein.

Cell Growth—Transformed E. coli cells were grown to log phase (A600 = 0.5) at 28°C in the presence of ampicillin (50 μg/ml) and chloramphenicol (50 μg/ml). Cultures were then induced with 0.2 mM isopropyl β-D-thiogalactoside (IPTG) and harvested by centrifugation. Cells from 20 liters of culture were resuspended in 200 ml of lysis buffer (20 mM NaP04, pH 7.0, 1 mM β-mercaptoethanol, 1 mM EDTA, and 1500 μM PMSF) and stored at -80°C.

Purification—All procedures were carried out at 4°C. The cell lysate from 20-liter culture was harvested, diluted in 2.5 volumes of 20 mM Tris-HCl, pH 8.5, 100 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol, 1 mM PMSF, and stored at -80°C.

Partial Proteolytic Digestions—p56aa and p110RB in S-200 buffer

RESULTS

PURIFICATION OF RB PROTEINS—Two species of RB were expressed from the pET8c construct, the full-length protein, p110RB, and p56aa, an amino-terminal truncated form of the protein. The amino terminus of p56aa was determined by microsequence analysis and found to correspond to Met-379 of the full-length protein. Expression of p56aa from Met-379 occurred due to the presence of an internal initiation site in the full-length RB cDNA. p56aa was the major product expressed, with the full-length protein only constituting approximately 10% of the total RB produced. The reason for the low yield of p110RB is not understood. p110RB was resolved from p56aa during the final purification step, gel filtration on a Sephacyrl S-200 column (Fig. 1A). RB proteins were purified to near homogeneity by SDS-PAGE (Fig. 1B). Following this purification protocol, the yields from a 20-liter culture were typically 2 mg of p110RB and 15–20 mg of p56aa. Both proteins were capable of binding to SV40 large T antigen and caused cell growth arrest in early G1 when microinjected.

2 T. Durfee, M. Mancini, D. Jones, S. J. Elledge, and W. H. Lee, submitted for publication.
into the osteosarcoma cell line, Saos-2, indicating that the purified proteins were biologically active (Goodrich and Lee, 1993a).

Partial Tryptic Digestion of p56RB—To explore the structural domain(s) of the RB protein biochemically, p56RB was incubated with trypsin as described under "Materials and Methods," and the progress of proteolysis was followed by analyzing samples from various time points on SDS-PAGE (Fig. 2A). p56RB was first cleaved to a transient 48-kDa fragment (Fig. 2, lane 2), which was further cleaved to two protease-resistant fragments, labeled A and B, of molecular mass 24 and 19 kDa, respectively (Fig. 2, lanes 3–7). Both proteolytic products A and B were resistant to further digestion for up to 2 h under the conditions described, even at a 10-fold higher ratio of trypsin to p56RB (data not shown). It is evident from these results that the sites of cleavage of p56RB, which contains a total of 78 potential trypsin cleavage sites scattered throughout the molecule, are restricted to a small segment of the molecule indicating a structural hierarchy in the protein causing many potential sites to be inaccessible.

![Figure 1. Purification of p110RB and p56RB.](image1)

Initial identification of these proteolytic products was performed by immunoblot analysis with specific antibodies against known regions of the RB protein. As shown in Fig. 2B, the monoclonal antibody against the A domain, α-1E5, recognized the 24-kDa protease-resistant fragment (Fig. 2B, lanes 3–7), and an antibody against the B domain, α-B, recognized the 19-kDa fragment (data not shown). Thus, the genetically defined A and B domains of the T-antigen binding region correspond to these two protease-resistant fragments, A and B, of the protein.

Amino-terminal Sequence Analysis Defines the Boundaries of the Proteolytic Fragments—To define further the precise amino-terminal boundaries of these domains, the proteolytic products, resolved from each other by SDS-PAGE, were subjected to automated Edman degradation. The analysis revealed that the 24-kDa fragment had an amino-terminal sequence of MNTIQQLM, and the 19-kDa fragment had an amino-terminal sequence of VNSTANAE, which correspond to p110RB residues 379–386 and 622–629, respectively (Fig. 5A). The first methionine in domain A is the first amino acid in p56RB indicating that p56RB is not digested at all from the amino terminus. Initial digestion is occurring from the carboxyl terminus, which is highly susceptible to proteolytic degradation as evidenced by the fact that the carboxyl-terminal antibody α-11D7 did not recognize any of the proteolytic fragments generated (data not shown). The resulting transient 48-kDa fragment is then cleaved in two, yielding two protease-resistant fragments.

These domains contain a total of 8 cysteine residues. A naturally occurring mutation affecting one of these cysteines is known to eliminate RB's T antigen and DNA binding activity. Cysteines commonly contribute to structural interactions via hydrogen bonds, disulfide bonds, or coordinate metal ion binding. To check whether these structural domains were linked by disulfide bond formation the proteolytic digest was run on a nonreducing gel. The migration pattern of the bands was found to be the same as under reducing conditions (data not shown) indicating that the domains were not linked by disulfides. This observation is supported by a recent report, which showed that 4 cysteine residues within the binding pocket were necessary for binding to RB-associated proteins, but these residues do not seem to be involved in the formation of disulfide bonds (Stirdivant et al., 1992).

Trypsin-resistant Fragments of RB Are Resistant to Other Proteases—To confirm that the discrete structural domains...
were not unique to trypsin digestion, p56RB was also subjected to cleavage by chymotrypsin and subtilisin, proteases that differ in their substrate specificities compared with trypsin. Digestion of p56RB with both these proteases produced a similar digestion pattern and yielded two major proteolytic fragments similar to those obtained by trypptic digestion (Fig. 3). In the presence of chymotrypsin p56RB was first cleaved to a transient 45-kDa fragment (Fig. 3A, lanes 1 and 2), which was in turn cleaved to yield two fragments of 24 and 17.5-kDa (Fig. 3A, lanes 3, 4, and 5). These correspond to fragments A and B from the trypptic digestion (as determined by immunoblot analysis) except that the B domain is slightly smaller and hence is labeled B'. Digestion by subtilisin proceeded in a similar fashion with a 46-kDa intermediate (Fig. 3B, lane 1) being cleaved in two to give two fragments of 24 and 19-kDa, which are labeled A and B, respectively (Fig. 3B, lanes 3, 4, and 5). As with the trypptic digestion the fragments produced by chymotryptic digestion were very stable to further digestion even at a 5-fold increase in the ratio of protease to RB. Subtilisin, however, could totally digest the protein after longer incubation times and at higher amounts of proteolytic enzyme. This is to be expected because subtilisin cleaves peptide bonds nonspecifically. The fact that three proteases of differing specificity cleave p56RB in close proximity to each other indicates that something other than the specificity of the protease is governing proteolysis.

**Partial Tryptic Digestion of p110RB**—To test whether the full-length protein, p110RB, contained the same and/or additional structural domains, p110RB was also subjected to trypptic digestion. Digestion of p110RB also occurred in a stepwise fashion with transient fragments giving way to more stable proteolysis-resistant domains (Fig. 4A). Tryptic digestion of p110RB yielded four protease-resistant fragments (Fig. 4A, lane 4) labeled A, B, N, and R of molecular mass 24, 19.5, 30, and 10 kDa, respectively. Due to its small size, the R fragment can be seen clearly in the original gel but poorly after photography. The amino terminus of each peptide was sequenced to identify the exact location of these domains in the full-length protein. The amino-terminal sequence of A was found to be TVMNTIQQLM, which corresponds to amino acids 377-386 in p110RB. Therefore, this is the same A domain as found in p56RB with cleavage occurring two amino acids before the first methionine of p56RB. The amino-terminal sequence of B was found to be identical to that of domain B obtained from the p66RB digestion.

The 30-kDa fragment, labeled N, was first identified by the amino-terminal anti-peptide antibody, α-D (Fig. 4B), which recognizes RB amino acids 62-91, indicating that this fragment originates from the amino terminus of RB. Amino-terminal sequence analysis of the 30-kDa domain revealed the residues ITAATAAAA, which correspond to amino acids 8-17 in p110RB. Fragment R was found to have the amino-terminal sequence IALQLENDTN, corresponding to amino acids 263-270 in p110RB (Fig. 5). A similar pattern was also obtained upon digestion of p110RB with subtilisin (data not shown) indicating that some structural organization in the protein is influencing proteolytic digestion. Whether R is a distinct domain or a sub-domain of N cannot be completely determined. However, the production of R corresponds to the formation of N from its higher molecular weight precursor, suggesting that there is a susceptible site between these two regions.

**Oligomerization of p110RB Revealed by Nondenaturating Gel Electrophoresis**—Analysis of the purified proteins revealed an interesting biochemical property specific to the full-length RB protein. Following nondenaturing polyacrylamide gel electrophoresis multiple bands were revealed upon immunoblot analysis using an anti-RB antibody (Fig. 6A, lane 2). At least seven bands corresponding to various oligomers of the protein were detected. Because the distance migrated by a protein on a native gel is a complicated function of both protein size and charge it is difficult to establish the exact molecular weight of the oligomers. The ladder-like appearance of the bands observed suggested that each species may differ from the adjacent one by unit mass. Plotting distance migrated by each band in Fig. 6A, lane 2, against log10 (n), where n equals the ladder position from the bottom, we found a nearly linear relationship (Fig. 6B). This supports the notion that each band differs from the next one by the addition of another RB molecule. Recombinant RB from a baculovirus expression system, purified in a different manner, also produced a similar pattern (data not shown), eliminating the possibility that this may be an artifact of a purification protocol. When an amino-terminal truncated form of the protein, p56RB (amino acids

---

**Fig. 3.** Time course of proteolysis of p56RB by chymotrypsin and subtilisin. A, purified p56RB (72 μg) was digested with chymotrypsin (144 ng, 66 units/mg) in 150 μl of buffer, 20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1 mM EDTA, and 10 mM 2-mercaptoethanol at 37°C. Aliquots (15 μl) were taken at the indicated time points and analyzed by SDS-PAGE (12%). B, purified p56RB (48 μg) was digested with subtilisin (48 ng) in 40 μl of buffer, 20 mM sodium phosphate, pH 7.5, 200 mM NaCl, 1 mM EDTA, and 10 mM 2-mercaptoethanol at 37°C.
Structural Domains of the p110RB

379–928) (Fig. 6A, lane 3), was analyzed by a similar method, only a single band was detected (Fig. 6A; lane 4). Pretreatment of full-length RB protein with dithiothreitol (30 mM) failed to abolish the formation of multiple bands, suggesting that oligomerization of the RB protein is mediated through an interaction other than disulfide bond formation (data not shown). The multiple band pattern was enhanced in the presence of a small amount of SDS (0.002%), suggesting a slight conformational change such as that induced by post-translational modification or binding to another protein might be necessary for this process to occur in vivo.

Electron Microscopy of RB Shows the Oligomeric Structure—When the retinoblastoma protein was analyzed by electron microscopy linearly extended macromolecular structures were observed (Fig. 7). The protein was placed directly on carbon/Formvar-coated copper grids, negatively stained with uranyl acetate, and viewed with an electron microscope. Immunogold labeling using an anti-RB antibody, Ab 245, shows numerous gold particles bound often in an apparently undulating or spiral manner along the structures indicating that these structures contain the RB protein (Fig. 7, a–c). When the experiment was repeated without adding the RB protein, the labeled structures did not appear, indicating that the structures do not originate from any of the reagents used in the immunogold-labeling experiment. The linearly extended structures were not detected when purified p56RB was examined (data not shown). Neither an anti-p53 antibody, Ab122 (Fig. 7d), nor an antibody directed against SV40 large T antigen, Ab419 (data not shown), labeled the structures eliminating the possibility that the structures seen were due to nonspecific antibody binding. These data suggest that the observed structures consisted mainly of the RB protein, providing another indication that the retinoblastoma protein forms macromolecular structures.

Intermolecular Interactions of the RB Protein—To identify regions of the retinoblastoma protein that could interact with each other the yeast two-hybrid system, which has proven to be a sensitive method for detecting protein-protein interactions under physiological conditions (Fields and Song, 1989), was used. In this system the ability of two fusion proteins to interact in vivo leads to expression of a lac Z reporter gene. Quantitation of β-galactosidase activity then serves as a reflection of the strength of the interaction occurring at the promoter. It has been demonstrated that the physical association detected in this system reflects the known in vitro binding properties of the RB protein and SV40 large T antigen (Durfee et al., 1993). Using this assay we have studied RB's intermolecular interactions and identified the domains that interact with each other.

Two fusion proteins, one containing the carboxyl-terminal half of RB (AA 301–928) fused to the Gal4-transactivating domain and a second containing the amino-terminal half of RB (AA 1–300) fused to the Gal4 DNA binding domain, were constructed as described under “Materials and Methods.” The yeast strain, Y153, was cotransformed with the plasmids encoding these fusion proteins. Cotransformants were assayed for their ability to activate transcription of the lacZ gene, and

---

**Fig. 4.** Time course of proteolysis of p110RB by trypsin. Purified p110RB (76 µg) was digested with 76 ng of TPCK-treated trypsin in 400 µl of buffer as described for p56RB. At various time points 40 µl of the reaction mixture was removed and the reaction stopped by the addition of PMSF (40 mM). A, analysis of proteolytic samples by SDS-PAGE (12%). N, A, B, and R indicate the proteolytic domains described in the text. B, immunoblot analysis of tryptic digest with α-D.

**Fig. 5.** Schematic representation of the structural domains in the retinoblastoma protein. The structural domains of p56RB (A) and p110RB (B) as determined by partial tryptic digestion and amino-terminal sequence analysis are shown. N, A, B, and R indicate the proteolytic domains described in the text.
by immunoblotting using Ab245, an anti-RB antibody expressed in protein (0.8 pg) (in 50 mM NaPO₃, pH 7.5, 250 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol, and 10% glycerol) was electrophoresed through a 6% polyacrylamide nondenaturing gel and analyzed described for and stained with Coomassie Blue. B, distance migrated by each band in the resulting β-galactosidase activity was measured. Analysis shows that the amino and carboxyl termini of RB, expressed in E. coli, was electrophoresed through a 7.5% SDS-PAGE and stained with Coomassie Blue (lane 1). Purified full-length RB protein (0.8 µg) (in 50 mM NaPO₃, pH 7.5, 250 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol, and 10% glycerol) was electrophoresed through a 6% polyacrylamide nondenaturing gel and analyzed by immunoblotting using Ab245, an anti-RB antibody (lane 2). Arrowhead denotes the position where the RB protein was loaded. Purified p56 faithfully expressed in E. coli, was analyzed as described for lane 1 (lane 3). Purified p56 protein (0.5 µg) (in 50 mM NaPO₃, pH 7.5, 250 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol, and 10% glycerol) was electrophoresed as described in lane 2 and silver stained (lane 4). Molecular size markers are in kilodaltons. B, distance migrated by each band in lane 2 of A was plotted against log10 (n), where n equals the band position from the bottom.

Purified ~56 protein (1.5 pg) (in 50 mM NaPO₃, pH 7.5, 250 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol, and 10% glycerol) was electrophoresed as described in lane 2 and silver stained (lane 4). Molecular size markers are in kilodaltons. B, distance migrated by each band in lane 2 of A was plotted against log10 (n), where n equals the band position from the bottom.

the resulting β-galactosidase activity was measured. Analysis shows that the amino and carboxyl termini of RB protein can interact as evidenced by a 34-fold increase in transcription (Fig. 8D). To determine the effect of mutations in the amino-terminal region on this interaction a number of fusion proteins containing deletions or insertions in this region were also constructed and tested in this assay. All mutations, except for an insertion at amino acid 181, abolished binding, indicating the importance of the amino-terminal region in this interaction. The region examined contains the two structural domains mapped to the amino terminus of RB, and all deletions within this region have deleterious effects on interactions between the amino and carboxyl termini. This indicates the importance of the structural integrity of this region for the N-C interaction to occur. The carboxyl-terminal end of RB did not have the ability to self-associate as no interaction was detected upon co-transfection of two fusion proteins, which each contained the carboxyl-terminal end of RB (Fig. 8B).

**FIG. 7.** Electron microscopy of purified RB protein. Immunogold labeling was performed on structures observed in purified full-length RB protein using Ab245 (0.10 mg/ml) (a-c) and Ab122 (0.11 mg/ml) (d) as the first antibody. For the second antibody, anti-mouse IgG goat antibody conjugated to 5-nm gold particles (Amersham Corp.) was used. Bar, 133 nm (a-c); 175 nm (d).

**DISCUSSION**

To date there are a number of reports on the purification of recombinant RB from E. coli and baculovirus systems (Edwards et al., 1992; Huang et al., 1991; Wang et al., 1990a). While truncated forms of the RB protein have been produced, the successful production of the full-length retinoblastoma protein has proved difficult. We have previously produced p110 from a baculovirus system, but yields were poor, and the protein tended to be denatured. The above purification scheme allows the isolation of significant amounts of stable and functionally active p110 suitable for biochemical studies.

Analysis of the purified proteins by partial proteolytic digestion revealed a structural organization in the protein previously unidentified. p56 is initially digested from the carboxyl terminus, which is likely an extended region of polypeptide highly susceptible to proteolytic cleavage. The resulting transient 45–48-kDa fragment is then cleaved in two, yielding the highly ordered structural domains A and B that correspond to the two regions previously defined as being required for binding to SV40 large T antigen. The domains mapped by deletion analysis fall within the two structural domains defined by partial proteolytic digestion. The exact carboxyl termini of these structural domains have not been determined, but they at least extend to the boundaries defined by deletion analysis, as judged by the molecular mass of the proteolytic fragments. Interestingly, these discrete domains do overlap with the positions of naturally occurring mutations in RB. The extreme susceptibility of RB to proteolytic cleavage in a region linking A and B strongly suggests that the continuity of the two domains is maintained by an accessible hinge region.

Using the Chou-Fasman method (Chou and Fasman, 1974) for predicting α-helices, β-sheets, and turns and the Karplus-Schultz method (Karplus and Schultz, 1985) for predicting the flexibility of the polypeptide chain the secondary structure of RB was analyzed. Domains A and B were found to have a much more defined secondary structure, while the region...
A.

Structural Domains of the p110RB

**FIG. 8.** Interaction between the amino and carboxyl termini of RB. A, fusion constructs used to test the interactions of the amino-terminal end of RB with its carboxyl-terminal end in yeast. The N-RB fusion proteins contained the Gal4 DNA binding domain joined to the amino-terminal fragment of RB, amino acids 1-300, in the pAS1 expression vector, creating pAS/N-RB. The Gal4 transactivation domain was fused to the carboxyl-terminal end of RB, amino acids 301-928, in the pJBTB expression vector, creating pJD1. B, determination of β-gal activity in transformed yeast cells. Y153 was transformed with various plasmids as indicated, and β-galactosidase activity was determined by the colony lift method and quantitated using an ONPG assay.

**TABLE 1.** Colony Fold increase in transcription

| Plasmid Name          | Colony Color | Fold Increase |
|-----------------------|--------------|--------------|
| None                  | white        | 1            |
| N-RB (1-300)          | blue         | 34           |
| N-HE (del 76-181)     | white        | <2           |
| N-BP (1-246)          | white        | <2           |
| N-BE (1-181)          | white        | <2           |
| N-HR (76-300)         | white        | <2           |
| N-HBg (ins 76)        | white        | <2           |
| N-EBg (ins 181)       | blue         | 46           |
| Pvu-C (246-928)       | white        | <2           |

"hinge" between the two domains was predicted to be a highly flexible turn region. Interestingly, this hinge region is also proline-rich compared with domains A and B. The carboxyl-terminal "tail" was the region with the least defined secondary structure and was predicted to be highly flexible with many turns. Nonetheless, this region is still important functionally, for example it contains the DNA binding domain (Wang et al., 1990a). It seems that p56RB is digested by various proteases in the flexible regions of least defined secondary structure. This is consistent with other reports, which have indicated that cleavage sites in proteins can be associated with flexible hinge regions, protruding surface loops, or similarly exposed segments lacking a defined secondary structure (Fontana et al., 1986).

A structural analysis of p60RB, an amino-terminal truncated RB that can be considered equivalent to p56RB, has been reported (Edwards et al., 1992). The ultraviolet circular dichroism spectrum of purified p60RB suggested that the protein contained 30% unordered conformations. Analysis of p56RB by partial proteolytic digestion also suggests that approximately 30% of the protein contains unordered conformations that are presumably very susceptible to proteolytic attack with the remaining 70% forming two highly ordered structural domains within the protein.

p110RB also contained the two structural domains, A and B, which seem to be distinct structural entities within both the full-length protein and the amino-terminal truncated form. In addition, this analysis led to the identification of two structural domains in the amino-terminal portion of the retinoblastoma protein, a large 30-kDa domain and a small 10-kDa domain. Given that the carboxyl-terminal half of RB is biologically active in growth suppression, the question remains as to the function of the amino-terminal half of the protein. The identification of two structural domains in the poorly studied amino-terminal end of RB may be indicative but is by no means proof that this region of the protein may have some biological function.
In this regard, the oligomerization of purified p110RB, which requires the amino-terminal domains, provides a new basis for conceptualizing its role inside the nucleus. The gel migration pattern obtained is indicative of proteins with oligomerization properties such as SV40 large T antigen (Prives et al., 1991) and several DNA helicases (Patael and Hingorani, 1993). These results suggest that the oligomerization is an intrinsic property of the full-length RB protein, requiring the presence of the amino-terminal region. Results from the yeast two-hybrid system suggest that there is a direct interaction between the amino and carboxyl termini of the protein, which would allow the formation of an oligomeric structure. Of course, the possibility of intramolecular interactions also exists with the amino and carboxyl terminus of the same molecule interacting. Interestingly, the other tumor suppressor gene, p53, can also oligomerize (Kraiss et al., 1988; O'Reilly and Miller, 1988). It has been reported that the unphosphorylated RB protein binds to several cellular proteins, including a number of transcription factors, and that such interactions may be disrupted when the RB protein is phosphorylated in late G1, S and M phases (Chellappan et al., 1991; Shirodkar et al., 1992). Though speculative, perhaps the RB protein may serve to regulate the activities of these other proteins by forming a "corral" that sequesters these factors (Lee et al., 1991). Whether oligomerization of the RB protein is a cell-cycle-regulated phenomenon, occurring in vivo, is currently being investigated.

For the first time it has become possible to examine at the structural level the retinoblastoma gene product. These studies provided new insights into the structural organization of the retinoblastoma protein revealing the presence of four distinct structural domains in the protein linked by exposed hinge regions. Some functions have already been defined for the carboxyl-terminal domains, and the identification of discrete structural domains in the amino terminus suggests this region may have specific biological function(s), which await further exploration. A novel property associated with the amino terminus, oligomerization, will provide an additional insight into the structure/function relationships of this tumor suppressor gene product and perhaps give some clues as to its mechanism of action in cell growth regulation.

Acknowledgments—We thank Dr. J.-S. Huang, St. Louis University, and Dr. P. Riflemen, University of Texas Health Science Center at San Antonio, for amino-terminal sequence analysis, C. M. Chang for providing assistance with electron microscopy, W.-M. Yang for technical assistance, R. Bookstein, R. E. Hollingsworth, and D. Sharp for helpful comments regarding the manuscript, and Pascale Rio for suggestions.

REFERENCES

DeCaprio, J. A.; Ludlow, J. W.; Figg, J.; Shew, J.-Y.; Huang, C.-M.; Lee, W.-H.; Mardusillo, E.; Paucha, E.; and Livingston, D. M. (1988) Cell 54, 275-286.
DeCaprio, J. A.; Ludlow, J. W.; Lynch, D.; Furukawa, Y.; Griffin, J.; Panaci-Worms, H.; Huang, C.-M.; and Livingston, D. M. (1989) Cell 58, 1085-1095.
DeCaprio-Jones, D.; Huang, P. S.; Jones, R. E.; Haskell, K. M.; Vuocolo, G. A.; Hanobik, M. G.; Huber, H. E.; and Oliff, A. (1991) Nature 352, 251-254.
Duree, T.; Becherer, K.; Chen, P.-L.; Yeh, S.-H.; Yang, Y.; Kilburn, A. E.; Lee, W.-H.; and Heimbrook, D. C. (1992) J. Biol. Chem. 267, 7971-7974.
Edwards, G. M.; Huber, H. E.; DeFeo-Jones, D.; Vuocolo, G.; Goodhart, P. J.; Magi, R. Z.; Sanyal, G.; Oliff, A.; and Heimbrook, D. C. (1992) J. Biol. Chem. 267, 7971-7974.
Fields, S., and Song, O. K. (1989) Nature 340, 245-246.
Fontana, A.; Fassina, G.; Viti, C.; Dalloppo, D.; Zami, M., and Zambonin, M. (1986) Biochemistry 25, 1547-1555.
Friend, S. H.; Bernards, R.; Rogelj, J.; O’Reilly, R., and Lee, W.-H. (1992) Science 252, 643-647.
Goodrich, D., and Lee, W.-H. (1993a) Nature 360, 177-179.
Goodrich, D., and Lee, W.-H. (1993b) Biochim. Biophys. Acta 1155, 43-61.
Goodrow, D., and Wang, N.-P. (1987) Science 232, 642-645.
Hollingsworth, R. E., Hensey, C. E., and Lee, W.-H. (1993) Curr. Opin. Genet. & Dev. 3, 55-62.
Huang, Y.; Park, S.-H.; Bogervann, E.; Cheng, J.-C.; Yandell, D. W.; Kave, J. F.; Minna, J. D.; Dryja, T. P., and Weinberg, R. A. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 2775-2779.
Huang, Y.-H.; Yee, J.-K.; Chittenden, T.; Li, Y.; Famham, P. J.; Blanar, M. A.; Livingston, D. M.; and Fleming, E. K. (1992) Cell 70, 351-364.
Karpfis, P. A., and Schultz, G. E. (1981) Naturwissenschaften 72, 5212-5217.
Kreider, R., Quiser, A., Dren, M., and Montenathan, M. (1988) J. Virol. 62, 4737-4744.
Laemmli, U. K. (1970) Nature 227, 680-685.
Leitch, J., and Lee, E. Y.-H. P. (1991) Nature 352, 64-65.
Lee, W.-H.; Shew, J.-Y.; Fang, F.; Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.
Lee, W.-H.; Bookstein, R.; Chiang, W., Fesick, J. M.; and Lee, E. Y.-H. P. (1993) J. Biol. Chem. 268, 14846-14851.