Isolation and Characterization of Complementary DNA to Proliferating Cell Nucleolar Antigen P40

A. Bramham Reddy, Amitava Chatterjee, Lawrence I. Rothblum, Ann Black, and Harris Busch

Department of Pharmacology, Baylor College of Medicine, Houston, Texas 77030

ABSTRACT

Proliferating cell nucleolar antigen P40 is a late G1-specific protein, which was found in a variety of human tumors (A. Chatterjee, J. W. Freeman, and H. Busch. Cancer Res., 47: 1123-1129, 1987). Two overlapping complementary DNA clones for antigen P40 were isolated by immunoscreening a λgt11 human expression library. The complete nucleotide sequence of the clones was determined. The complementary DNAs encode the M, 30,000 portion of the COOH-terminal portion of the protein. The mRNA for P40 was 2.8 kilobases long and was expressed maximally in G1 cells in cell cycle. A series of deletion mutants of the expressed peptide was constructed and the deletion mutants were expressed in Escherichia coli. Using these mutants, the epitope region of P40 recognized by a P40-specific monoclonal antibody was identified. The hydrophathy plot based on the protein sequence revealed that this region of the protein is largely hydrophilic. This protein is unique and differs in sequence from other proliferating cell nuclear/nucleolar antigen proteins of similar molecular weight such as protein B23 and cyclin.

INTRODUCTION

The nucleoli of tumor cells are morphologically distinct from nontumorous cells (1); this has been attributed to a very high level of nucleolar activity such as rRNA synthesis and the assembly of preribosomal particles. Biochemical (2-4) and immunological (5-7) differences between the nucleoli of tumor and normal cells have been reported. Immunological analyses have detected PCNA in actively dividing cells and in tumor tissues, but not in the normal (resting) cells (8-11). The precise roles of these proteins in cell proliferation, in transformation, or in the maintenance of the malignant phenotypes are not yet known. Initial suggestions have been made that they may be involved in cell growth (6, 7, 9), in expression of oncofetal proteins (4), or in transcriptional control (12).

Monoclonal antibodies specific for nucleolar proteins have been developed in our laboratory (13-16). During these studies, a novel M, 40,000 nucleolar antigen (P40) was identified. Monoclonal antibodies for P40 were obtained when a HeLa nucleolar protein extract, separated from the immunodominant protein C23 was used to immunize mice (13). P40 was not detected in normal liver, kidney, and leukocytes but was readily demonstrable in a variety of human malignant tissues. Characterization of two proteins of similar molecular weight, B23 (17) and cyclin (18), has been reported recently. Cyclin, which has a molecular weight of 35,000, was the first PCNA which was found to be associated with G1-S phase of the cell cycle (9). Cyclin has been shown to be an auxiliary protein of DNA polymerase δ (19). P40 differs in its specific nucleolar localization from cyclin, which is largely in the nucleoplasm. In addition, cyclin appears in the nucleolus in S phase; P40 appears in the nucleolus in G1, 6 h after refeeding serum-starved HeLa cells.

The present study was initiated to isolate cDNA for antigen P40 and to determine whether structural relationships exist among P40, B23, and cyclin, which have similar molecular weights. The cDNA would also be useful for the structural and functional analyses of P40 and as a probe to study its expression in various tissues under various conditions. Another goal of this study was to identify the epitope region for P40-specific monoclonal antibody on the P40 protein sequence. Epitopes present on the PCNA proteins would serve as possible targets for anticancer drugs (20), and information on the amino acid sequence of the epitope is essential for development of useful therapeutic information.

The amino acid sequence derived from the cDNA clones isolated in this study shows that P40 is a unique protein and differs from previously described PCNAs, protein B23 and cyclin. The binding region (epitope) for a P40-specific monoclonal antibody has been localized.

MATERIALS AND METHODS

DNA-modifying enzymes were purchased from Boehringer Mannheim or Bethesda Research Laboratories and used according to their instructions. Plasmid vector pGEM-3Z and protoblot immunoscreening system were obtained from Promega Biotech. Hep-2 cDNA library in λgt11 was kindly supplied by Dr. Harry Harris (University of Pennsylvania). The monoclonal antibodies for protein P40 were characterized previously (13). [a-32P]dATP was purchased from Amersham. cDNA Library Screening. The Hep-2 cDNA library in λgt11 was screened with a monoclonal antibody (MAb-15) for P40 (21). Approximately 1 x 10^6 phage plaques were plated on Escherichia coli Y1090 and the expression of fusion proteins was induced with IPTG. The proteins were transferred to nitrocellulose filters. The nitrocellulose filters were treated successively with P40-specific monoclonal antibody 15 and anti-mouse IgG conjugated to alkaline phosphatase. Positive clones for P40 were identified by subsequent color development with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate.

DNA Subcloning and Sequencing. The positive clones were plaque purified and the phage was grown in large volumes. The DNA was isolated from the recombinant clones and digested with EcoRI, and the inserts were purified from a preparative agarose gel. These inserts were subcloned in plasmid pGEM-3Z. A series of deletion mutants were constructed from the cDNA inserts by the use of exonuclease III and S1 and anti-mouse IgG conjugated to alkaline phosphatase. Positive clones for P40 were identified by subsequent color development with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate.

DNA Subcloning and Sequencing. The positive clones were plaque purified and the phage was grown in large volumes. The DNA was isolated from the recombinant clones and digested with EcoRI, and the inserts were purified from a preparative agarose gel. These inserts were subcloned in plasmid pGEM-3Z. A series of deletion mutants were constructed from the cDNA inserts by the use of exonuclease III and S1. 1

Received 7/5/88; revised 12/1/88; accepted 1/6/89.

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

1 These studies were supported by the Cancer Research Center Grant CA-10893, PI, awarded by National Cancer Institute, Department of Health and Human Services; the DeBakey Medical Foundation; The Davidson Fund; The Pauline Sterne Wolff Memorial Foundation; H. Leland Kaplan Cancer Research Endowment; Linda and Ronny Finger Cancer Research Endowment Fund; and the William S. Farish Fund.

2 To whom requests for reprints should be addressed.

3 The abbreviations used are: PCNA, proliferating cell nuclear/nucleolar antigen; IPTG, isopropyl-β-D-thiogalactopyranoside; cDNA, complementary DNA; SDS, sodium dodecyl sulfate; poly(A), polyadenylate; SSC, standard saline citrate (1 x SSC = 0.15 M NaCl-0.15 M sodium citrate, pH 7.0); MAb-15, monoclonal antibody 15.

1763
lulose column chromatography (27). RNA was subjected to electrophoresis on 1% agarose gels containing formaldehyde (28) and blotted onto NYTRAN paper (Amersham Inc.). For hybridization, the probes were labeled with the multiprime labeling system (Amersham Inc.) (29). Hybridization was carried out at 42°C in the presence of 50% formamide, 5 x SSC, 5 x Denhardt’s solution (1 x Denhardt’s = 0.02% each of Ficoll, polyvinylpyrrolidone, and bovine serum albumin), 0.1% SDS and 100 μg/ml denatured calf thymus DNA for 12 h. The filters were washed twice each for 20 min with 6 x SSC + 0.1% SDS at room temperature, 1 x SSC + 0.1% SDS at 37°C, and 1 x SSC + 0.1% SDS at 68°C. These filters were subjected to autoradiography.

Expression of Fusion Proteins in E. coli. Lysogens for the recombinant clones were isolated using E. coli Y1089. The overnight cultures of these lysogens grown at 32°C were brought to 42°C for 20 min, IPTG was added to a final concentration of 10 mM and incubation was continued for 1 h at 37°C on a shaker. Cells were pelleted and lysed by boiling in Laemmli buffer (24). The immunoreactive fusion proteins were analyzed by electrophoresing the E. coli extracts on SDS-polyacrylamide gels and by Western blotting.

For expression of fusion proteins from plasmid harboring E. coli, a similar protocol was followed. The reading frame of the insert DNA in plasmid was changed by the addition or deletion of a base in the multiple cloning site of the plasmid located 5' of the insert. A series of deletion mutants of P40 from the NH2 terminus were constructed by the use of exonuclease III and S1 nuclease (22). A few with the correct reading frame were selected and the fusion proteins coded by them were analyzed by Western blotting.

HeLa Cell Synchronization and Cell Cycle Expression of P40. HeLa cells were synchronized by a double thymidine block procedure (30). Following release from the second block, the cells were plated in 150-cm2 tissue culture flasks; 8 h later the mitotic cells (G2-M phase) were selectively detached. For the G1 time points, the mitotic cells were incubated for an additional 1.5 and 3.0 h at 37°C. The G1-S phase cells were obtained prior to release from the second block. The S-phase cells were collected 10 h following release of the second block.

Total RNA was prepared from these cells (26), and 40 μg of RNA were subjected to electrophoresis (28) and Northern hybridization using the P40.2 cDNA probe.

RESULTS

Approximately 1 x 10^6 phages of Hep-2 cDNA library in λgt11 were screened with the P40 monoclonal antibody 15. Two positive plaques were obtained, plaque purified, and designated P40.1 and P40.2. Lysogens for these two clones were isolated and induced with IPTG to express the fusion proteins. The E. coli extracts were electrophoresed and subjected to Western blotting (Fig. 1A). Both clones coded for an immunoreactive fusion protein with a molecular weight of approximately 150,000 (of which approximately 35 kD resulted from the inserts). When similar gels were stained with Coomassie Brilliant Blue R, strong bands corresponding to fusion proteins were observed in the IPTG(+)-lanes (data not shown). To confirm that these clones represented cDNAs for P40 and did not simply produce a common epitope, the fusion proteins were reacted with four noncompeting monoclonal antibodies specific for protein P40. Fig. 1B shows that all of these monoclonal antibodies reacted with the fusion proteins. In all cases, one major dense band and one minor band (possibly a product of proteolytic degradation) were found.

Sequence Analysis of P40 cDNA Clones. DNA was isolated from clones P40.1 and P40.2; their insert lengths were found to be 2.1 and 1.7 kilobases, respectively. These inserts were subcloned into plasmid pGEM-3Z. A partial restriction map showing the relationship between the two clones is shown in Fig. 2.

* A. Chatterjee and H. Busch, unpublished observations.

The nucleotide sequence of these cDNAs was determined using the strategy shown in Fig. 2. Clone P40.1 was 2066 nucleotides long (Fig. 3) and had an open reading frame of 792 base pairs corresponding to 264 amino acids (M, 30,000 protein). Clone P40.2 was 1706 base pairs long and contains nucleotides 10-1715 of clone P40.1. There were two polyadenylation signal sequences (AAATAAA) at 1541 and 1668 base pairs of the cDNA. The difference between the two clones is mainly at the 3' end. Since the open reading frame codes for only the M, 30,000 portion of the M, 40,000 protein, the cDNAs lack part of the NH2 terminus sequences and the 5' noncoding sequences. Northern blot analysis (Fig. 4A) also showed that the cDNAs do not represent the full length mRNA which is approximately 2.8 kilobases long.

Clone P40.1 has an open reading frame of 264 amino acids and the deduced amino acid sequence is shown in Fig. 3. This peptide is relatively rich in proline (7.6%). The hydrophilicity/hydrophobicity profile for this sequence calculated by the method of Hopp and Woods (31) shows the protein is largely hydrophilic (Fig. 5).

Expression of P40 mRNA. To study the regulation of P40 mRNA in the cell cycle, RNA isolated from different phases of the HeLa cell cycle was probed with P40.2 (Fig. 4B). P40-specific mRNA was maximally expressed in G1 cells (Lanes 2 and 3) which correlates G1-specific expression of the P40...
cDNA FOR THE P40 PCNA

CTG GCC CTC CCC ACC CTC ACC AGC GGC CTT TTC CGC TCT GAG TCC CTT GTC TCC CCI CUT GCC GAG CCG CTG CTG 75
Leu Ala Leu Pro Thr Leu Pro Arg Ser Ser Leu Glu Ser Pro Arg Gly Glu Arg Leu Leu 25
CAG SAT GCC ATC CTG CTA GAG GAT CTA AAA AAC CTC CTG GAG CCA GTA GAA GAA CTC GCT TGC ACG CCA 150
Gln Asp Ala Ile Arg Glu Val Glu Leu Lys Asp Leu Leu Leu Gly Pro Gly Val Leu Leu Thr Pro 50
CAA GAC CCA CTC CGG TTA GAA CCA AGC AAC GCT GTT ACT CAG TCA AGC TCT GCT AAT GCT GCT GGC 225
Arg Glu Pro Ala Leu Pro Leu Glu Glu Ala Arg Ala Arg Cys Ser Cys Ser Ser Thr Ser Ser Gly Ser 100
AAA ACC TTC AAT AAG GCT GCC CGC GTT ACG TGA ACT GAT CTA CCA CTT GTC AAT CTC TAT GGA CTT 290
Thr Arg Thr Pro Arg Gln Ala Gln Thr Ala Pro Gln Ser Leu Pro Ser Leu Tyr Gln Leu Gln Leu 150
AGA TCA CAG CCC TGT CCC ATT GCT TGT GGT TGC CTT GGG GTC ACT GCC AAT GGT GAG GCC 355
Arg Ser Pro Glu Cys Ser Pro Ser Ser Thr Leu Ser Ser Cys Leu Ser Ser Cys Ser Ser Gly Ser 200
GTG GCC GAG CAC GAG ACT CTG ATG GAA GCC CGC TCT CCT CCT CTA CCT GCC GAC CAG GCC CGC GCA CTGCTGGAG 420
Val Leu Glu Glu Arg Ser Leu ATG Glu Ala Glu Ala Ser Pro Ser Pro Ser Ser Ser Ser Glu Ser Ser Ser 250
GAG AAT TCG GCC AGT GCTGGC AGC CTGGAG GCCCGCCTCCCGGGAG AGT GAG CAG GCC CGC GCA CTGCTGGAG 495
Glu Thr TCA Ala Glu All Gly Ser Leu Gtu Ale Arg Leu Arg Glu Ser Glu Gtn Ale Arg Ate Leu Leu Leu Glu Arg Leu 300
GAT ACT CTC TCA GAG CAG CTG GCC GCC CTG GCC CAG ACC GAG CAA CTC CCA GCT GAG GCC CCC TGG GCC CGC AGA CCT GTG 570
Ser Thr Leu Ser Glu Gin Leu All Ala Leu Gty Gin Thr Glu Gin Leu Pro All Glu All Pro Trp All Arg Arg Pro Vll 350
CTG AGA TCA CCG CAA GAG GAG GCG TTA CAG CGA TTG GTC AAT CTC TAT GGA CTT CT> CAT GGC CTA CAG GCA GCT 645
Leu Arg Ser Pro Gin Glu Glu Ale Leu Gin Arg Leu Vet »11 Gly Leu Gin Ale Al« 400

Fig. 3. Complete nucleotide sequence of the cDNA insert P40.1. The nucleotide sequence of the coding strand and the corresponding predicted amino acid sequence are shown. The numbers on the right ordinate refer to nucleotides and amino acids.

Fig. 4. (A) Detection of P40 mRNA. Poly(A)-containing mRNA (2 µg) from HeLa cells was subjected to electrophoresis, transferred to NYTRAN paper, and hybridized with 32P-labeled P40.2 DNA. The mobilities of 28S and 18S rRNAs are indicated. (B) Cell cycle dependent expression of P40 mRNA. Forty µg of total RNA from different cell cycle stages of HeLa cells were electrophoresed and probed with 32P-labeled P40.2 probe. The different phases from which RNA was isolated are: Lane 1, G2-M; Lane 2, 1.5 h d; Lane 3, 3.0 h G,; Lane 4, G,-S; and Lane 5, S phase. Arrow, band corresponding to P40-specific mRNA.

protein (13). There was a low level of expression of P40 mRNA in G1-S cells (Lane 4) and no detectable amount in S phase cells (Lane 5).

Localization of the Epitope Region for MAb-15. To determine the essential region for the binding of monoclonal antibody (MAb-15) to protein P40, a series of mutants with various lengths of deletions from the NH2 terminus were constructed. The products expressed by these deletion mutants were analyzed by electrophoresis and Western blotting for their binding to MAB-15.

Fig. 6 shows that the deletion up to 240 base pairs did not affect the binding of the fusion protein to MAb-15. However, with deletions to nucleotide 276 and 351, the protein was no longer recognized by MAb-15, indicating that the region between 241 and 277 base pairs (Thr-Phe-Asn-Gly-Ser-Ile-Glu-Leu-Cys-Arg-Ala-Asp) is essential for binding of MAb-15 to P40. Hence this region forms part of the MAB-15 binding site and may represent the antigenic determinant. To show that the deletion peptides were P40 sequences, one construct was made with an incorrect reading frame (Fig. 6, P40.2-10”; Fig. 7) which was not producing any immunoreactive protein. In these deletion constructs, apart from the predicted size polypeptide (shown by the dot), there were two other larger size peptides from each of the constructs (Fig. 6) which probably represent translational readthrough products.
were detected by probing with MAb-15. The different deletion constructs used is indicated. P40.2-10" construct was not in correct reading frame. Predicted size fusion proteins. The presence and size of the immunoreactive fusion proteins 

tions in the coding region of P40.2 insert in plasmid pGEM-3Z were constructed./coli harboring these deletion plasmids was induced with IPTG to produce epitope region for a P40-specific monoclonal antibody. Two which have been sequenced.

The P40 sequence did not have any homology to nucleolin (32), cyclin (18), or protein B23 (17), three other nucleolar proteins were characterized by DNA sequence analysis. When the Gene OVERLAPPING DNA CLONES ENCODING THE PROTEIN P40 HAVE BEEN TO PROLIFERATING CELL NUCLEOLAR ANTIGEN P40 AND TO IDENTIFY THE EPITOME REGION FOR A MONOCLONAL ANTIBODY. A SIMILAR RATIONALE HAS BEEN USED TO MAP THE EPITOPES ON THE M, 190,000 GLYCOPROTEIN OF Plasmodium falciparum (34). BY ANALYZING FUSION PROTEINS FROM VARIOUS DELETION MUTANTS OF CLONE P40.2, THE REGION BETWEEN 241 AND 277 BASE PAIRS OF THE INSERT CODING FOR Thr–Phe–Asn–Gly–Ser–Ile–Glu–Leu–Cys–Arg–Ala–Asp, was found to be essential for the binding of MAB-15. Although this region of P40 is indispensable for MAB-15 binding, it may not represent the complete epitope sequence. Further, part of this 12-amino acid region could also be removed without affecting the antibody binding, and hence more mutants are required to define the exact NH2-terminal and COOH-terminal boundaries of the epitope. When this sequence is defined, the role of each amino acid can be analyzed by mutating them individually.

**REFERENCES**


16. Freeman, J. W., Busch, R. K., Gyorkey, F., Ross, B. E., and Busch, H. Identification and characterization of a human proliferation-associated nucleolar antigen with a molecular weight of 120,000 expressed of the coding region of the NH2 terminus were not present in these cDNA clones. There are two consensus polyadenylation signal sequences (AAATAA) at 1541 and 1668 base pairs of clone P40.1. In general, the polyadenylation signal sequence is present 10–15 base pairs upstream of the poly(A) stretch (33). Because the polyadenylation signal sequences from the 3' end in clone P40.1 are 398 and 525 base pairs upstream, and this clone does not have a poly(A) stretch, it is not certain if these two poly(A) signals are functional.

In the present study, we initiated analyses of localization of an epitope region on the P40 by constructing plasmids with deletions in the coding region and determining their capability to recognize the monoclonal antibody 15. A similar rationale has been used to map the epitopes on the M, 190,000 glycoprotein of Plasmodium falciparum (34). By analyzing fusion proteins from various deletion mutants of clone P40.2, the region between 241 and 277 base pairs of the insert coding for Thr–Phe–Asn–Gly–Ser–Ile–Glu–Leu–Cys–Arg–Ala–Asp, was found to be essential for the binding of MAB-15. Although this region of P40 is indispensable for MAB-15 binding, it may not represent the complete epitope sequence. Further, part of this 12-amino acid region could also be removed without affecting the antibody binding, and hence more mutants are required to define the exact NH2-terminal and COOH-terminal boundaries of the epitope. When this sequence is defined, the role of each amino acid can be analyzed by mutating them individually.

**DISCUSSION**

This study was undertaken to isolate complementary DNA to proliferating cell nucleolar antigen P40 and to identify the epitope region for a P40-specific monoclonal antibody. Two overlapping DNA clones encoding the protein P40 have been isolated by immunoscreening of a λgt11 expression library and were characterized by DNA sequence analysis. When the Gene Bank was screened with a Microgenie program for any similarity with P40, no homology was found at DNA or protein levels. The P40 sequence did not have any homology to nucleolin (32), cyclin (18), or protein B23 (17), three other nucleolar proteins which have been sequenced.

The mRNA size for P40, determined by Northern blot analysis, is 2.8 kilobases (Fig. 4A), which is longer than required to code for a M, 40,000 protein and suggests that there are long 5' and/or 3' noncoding regions. The cDNA clones obtained thus far lack portions of the 5' and 3' ends. Although clone P40.1 has a 1.4-kilobase untranslated region, it does not end with a poly(A) stretch. The 5' end untranslated region and part

3 P. K. Chan et al., submitted for publication.

---

**Fig. 6.** Localization of the epitope region for MAB-15 on protein P40. Deletions in the coding region of P40.2 insert in plasmid pGEM-3Z were constructed. E. coli harboring these deletion plasmids was induced with IPTG to produce fusion proteins. The presence and size of the immunoreactive fusion proteins were detected by probing with MAB-15. The different deletion constructs used is indicated. P40.2-10" construct was not in correct reading frame. *¿, predicted size peptide from each of the constructs; the bands above these could be arising from translational readthrough.

**Fig. 7.** Schematic representation of the results presented in Fig. 6.
cDNA FOR THE P40 PCNA


Isolation and Characterization of Complementary DNA to Proliferating Cell Nucleolar Antigen P40

A. Bramham Reddy, Amitava Chatterjee, Lawrence I. Rothblum, et al.


Updated version  Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/49/7/1763

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/49/7/1763. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.