Characterization of the Human p53 Gene

PETER LAMB AND LIONEL CRAWFORD*

Molecular Virology Laboratory, Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London, WC2A 3PX United Kingdom

Received 12 November 1985/Accepted 27 January 1986

Cosmid and λ clones containing the human p53 gene were isolated and characterized in detail. The gene is 20 kilobases (kb) long and has 11 exons, the first and second exons being separated by an intron of 10 kb. Restriction fragments upstream of sequences known to be within the first identified exon were tested for promoter activity by cloning them in front of the chloramphenicol acetyltransferase gene and transfecting the resulting constructs into HeLa cells. A 0.35-kb DNA fragment was identified that had promoter activity. Results of primer extension experiments indicated that the mRNA cap site falls within this fragment, as expected. Analysis of the sequence upstream of the presumptive cap site indicated that the human p53 promoter may be of an unusual type.

There is direct evidence that the cellular phosphoprotein p53 can contribute to transformation. Transfection of primary rodent cells with a cloned mouse p53 gene and an activated *ras* gene results in morphological transformation (13, 32), in a similar but not identical manner to that elicited by *myc* and *ras* or E1A and *ras* (19, 34). On the basis of this it has been suggested that p53 falls into the class of nuclear oncogenes; and indeed, under the control of a strong promoter, a mouse p53 cDNA clone can extend the life-span of primary cells in culture (18), an activity shared by *myc* and E1A (41).

Results of previous studies have demonstrated that elevated levels of p53 are a common feature of a wide variety of transformed cells of both murine and human origin (4, 10). Also, in cells transformed by simian virus 40 (SV40) or adenovirus, p53 forms a specific complex with either large T (20) or the E1B 58,000-molecular-weight protein (36), which are viral gene products known to be required for transformation.

The role of p53 in normal cells is unclear, although it has been shown that injection of anti-p53 antibodies prevents quiescent 3T3 cells from re-entering the cell cycle on serum stimulation (26, 27). This suggests the possible involvement of p53 in the G_0 to G_1 transition, an interpretation supported by the observation that mitogenic stimulation of lymphocytes with concanavalin A induces p53 synthesis (28).

The mouse genome contains two loci for p53, both of which have been cloned and characterized. One corresponds to an intronless pseudogene which probably arose via reverse transcription of p53 mRNA and subsequent reintegration into the genome (47). The other represents the functional gene which gives rise to an mRNA of 2.0 kilobases (kb). This gene is split into 11 exons and covers 12 kb of DNA (6).

In contrast, little is known about the arrangement of p53 gene(s) in the human genome. We have recently reported the isolation of a cDNA clone for human p53 (25). Using this partial cDNA clone as a probe, we identified only a few hybridizing bands on Southern blots of human DNA, which is consistent with the presence of one or a few p53 genes (25). Here we present the cloning of a cDNA that extends into the 5'-untranslated region of human p53 mRNA and the

isolation and detailed characterization of the human p53 gene.

MATERIALS AND METHODS

Construction of λ libraries of *Hind*III-cut human DNA. High-molecular-weight human fetal liver DNA was isolated (8), digested to completion with *Hind*III, and fractionated on an 0.8% low-melting-point agarose gel. Slices of the gel centered on 7 and 2.5 kb (as determined by coelectrophoresed markers) were removed and melted, and the DNA was extracted several times with 10 mM Tris hydrochloride (pH 8)–1 mM EDTA (pH 8)–saturated phenol and ethanol precipitated. The DNA was redissolved in 10 mM Tris hydrochloride (pH 8)–1 mM EDTA (pH 8) and ligated into a threefold molar excess of dephosphorylated *Hind*III-cut λ NM1149 (31) arms. The resulting recombinants were packaged in vitro and plated on *Escherichia coli* C600.

Isolation of clones from genomic and cDNA libraries. λ libraries were screened by the procedure of Benton and Davis (5), and the cosmid library was screened by the method of Hanahan and Meselson (16). The λ libraries of HindIII-cut human DNA were screened with the mouse p53 cDNA clone 9 under conditions of reduced stringency, as described previously (2). All other libraries were screened under high stringency conditions, defined as hybridization at 42°C in a solution containing 50% formamide, $6 \times$ SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) 100 µg of sheared salmon sperm DNA per ml, $5 \times D$ ($5 \times D$ is 0.1% bovine serum albumin, 0.1% polyvinylpyrrolidone, and 0.1% Ficoll), and 0.1% sodium dodecyl sulfate and washing at 60° C in 0.2× SSC and 0.1% sodium dodecyl sulfate. The human cDNA library was a gift from H. Kataoka and M. Collins, and the human cosmid library was a gift from S. Carson.

DNA sequencing. DNA fragments were subcloned into M13 vectors by standard techniques (24). Sequence determination was carried out by the dideoxy method (35). All sequences were determined on both strands, except for the sequence of exon 7, which was obtained from multiple readings of one strand.

Cloning p53 gene fragments into CAT vectors, eucaryotic cell transfection, and CAT assays. The 2.4-kb *EcoRI-XbaI* fragment and the 0.35-kb *BamHI-XbaI* fragment were isolated from low-melting-point agarose gels, purified on Elutip columns (Schleicher & Schull Inc., Keene, N.H.), and

^{*} Corresponding author.



FIG. 1. Map of the human p53 cosmid and λ genomic clones and the organization of the p53 gene. The following restriction enzyme sites are indicated: E, *Eco*RI; H, *Hind*III; B, *Bam*HI; X, *Xba*I; and S, *SaII*. Only the *Eco*RI sites are shown on cosp53PHD. The *SaII* sites are derived from the cosmid vector. *Bam*HI sites 5' to the one indicated within the 3.8-kb *Eco*RI fragment were not mapped. The positions of the subcloned probes HU2-6 and HU7-1 are indicated. The structure of the gene is shown on the lower line with the exons denoted by black boxes.

ethanol precipitated. The DNAs were redissolved in TE and blunt-ended with the Klenow fragment of DNA polymerase I. After phenol extraction and ethanol precipitation, the DNAs were ligated into an excess of dephosphorylated *SmaI*-cleaved pSVOsma. pSVOsma is a derivative of pSVO (14) that has a *SmaI* site in place of the *Hind*III site 5' to the chloramphenicol acetyltransferase (CAT) gene.

Plasmid DNAs to be used for transfection were purified twice on CsCl gradients. The day before transfection, HeLa cells (1) were seeded at a density of 5×10^5 cells per 90-mm-diameter dish in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum. The medium was replaced 4 h before transfection. Plasmid DNA (20 µg) was applied to each plate as a calcium phosphate precipitate (15). The precipitate was left on the cells overnight at 32°C, after which the cells were shocked with medium containing 15% glycerol for 1 min, rinsed with medium, and incubated at 37°C. After 48 h, the cells were harvested, and lysates were assayed for CAT activity (14). The growth state of the cells when harvested (e.g., whether confluent or exponentially dividing) did not affect the results.

RNA extraction and primer extension. RNA was isolated from SV80 cells (39) by the guanidinium isothiocyanate method (9). Poly(A)⁺ RNA was selected by passage over oligo(dT)-cellulose. A 100-fold molar excess of a synthetic oligonucleotide (Genetic Research Instrumentation) labeled at its 5' end with polynucleotide kinase (P-L Biochemicals, Inc., Milwaukee, Wis.) was annealed either to 5 µg of poly(A)⁺ RNA or to 5 µg of yeast tRNA (Sigma Chemical Co., St. Louis, Mo.) at 55°C in 5 µl of 0.4 M NaCl-0.01 M piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES). After 3 h, the hybridization mixtures were diluted into 80 µl of buffer containing 50 mM Tris (pH 8.2), 10 mM dithiothreitol, 6 mM MgCl₂, 25 µg of actinomycin D per ml, 0.5 mM deoxynucleoside triphosphates, and 10 U of reverse transcriptase (Anglian Biotechnology). The samples were incubated at 42°C for 1 h, and the reaction was terminated by heating to 70°C for 15 min. DNase-free RNase (Sigma) was added, and the incubation was continued at 37°C for 30 min. Finally, the samples were phenol extracted, ethanol precipitated, suspended in formamide dyes, and electrophoresed on 5% polyacrylamide gels containing 7 M urea.

RESULTS

Isolation of human p53 genomic and cDNA clones. Under conditions of reduced stringency, the mouse p53 cDNA clone 9 detects two hybridizing fragments of 2.5 and 7 kb on Southern blots of human DNA cut with HindIII (2). This cDNA was used to screen genomic libraries constructed by cloning HindIII digests of human DNA into phage λ . In this way clones corresponding to both the 2.5- and 7-kb bands were isolated. Nonrepetitive restriction fragments containing p53 sequences were subcloned from these phage to give plasmids HU2-6 and HU7-1 (Fig. 1). These plasmids have been used successfully to isolate a 2.1-kb human cDNA clone p102 (25), but because this clone lacked sequences corresponding to the N terminus of the protein, additional cDNA libraries were screened. A cDNA clone generated by using a synthetic oligonucleotide to prime cDNA synthesis on SV80 mRNA (G. Matlashewski, unpublished data) was used as a probe. p53J6K was isolated from a cDNA library constructed with mRNA from the T-cell line J6, a subclone of Jurkat (37). Sequence analysis showed that it was 503 base pairs (bp) long, of which 125 bp was 5'-untranslated region (Fig. 2). The sequence of p53J6K is in complete agreement with our previously published sequence for p102 (25) over the 78 bp of overlap between them, but differs by one nucleotide from the recently published sequences covering the 5' end of human mRNA. In place of the C residue at position 429 in the sequence of Harlow et al. (17), which is nucleotide 213 in the sequence of Zakut-Houri et al. (45), we found a G residue (Fig. 2). This results in an arginine codon (CGC) instead of a proline codon (CCC). The sequence of the genomic clone (see below) also contained a G at this position and is thus in agreement with the p53J6K sequence.

Southern blot analysis of human fetal liver DNA with a variety of probes indicated that there is a single gene for p53 in the human genome, the 5' end of which is on a 3.8-kb *Eco*RI fragment; the rest of the gene is on two *Eco*RI fragments of 16 kb. To obtain genomic clones spanning the

AGGTGTGGATATTACGGAAAGCCTTCCTAAAA 100 CENTCENCETENCING TCTTCCTTCCACCCTTC 150 200 TOCTOCANAATCAT CCACCCCAAAATC 300 TCTGCACCCTCCTCCTCC CACCOTCCAGGGAGCAGGTAGCT 550 **GTTCGGGCTGGGAGCGTGCTTTTCCACTA** Intron 1 10 000 bo TOCCTTOCOCOTCACT E E CAG TCA GAT OCT AGE GTE GAG COE COT CTG AGT CAG GAA ACA TTT Q S D P S V E P P L S Q E T P TCA GAC CTA TGG AAA CT --- Intron 2 117 bp --- A CTT CCT S D L W K L I. P 10 AAC AAC GTT CTG --- Intron 3 93 bp TCC CCC TTG CCG TCC CAA GCA ATG GAT GAT TTG ATG CTG TCC CCG S P L P S Q A N D D L N L S P 40 GAC GAT ATT GAA D D I Z N 1 Z TOG TTC ACT GAA GAC COA GGT Q TAC GOT TTC COT CTC Y G P R L 110 GOC TTC TTG CAT TCT GOG ACA GOC ANG TCT GTG ACT TGC ACG-- Intron 4 800 bp - -TAC TCC CCT GOC CTC ANG ANG ATG TTT TGC G P L H S G T A R S V T C T Y S P A L H K N P C 120 Y S P A L H K N P C 1 30 1000 1100 CAG CAC ATG AGG GAG GTT GTG AGG COC TOC COC CAC CAT GAG COC TOC TCA GAT AGC GAT G--- Intron 5 81 bp --- GT CTG GCC 0 8 M T E V V R R C P 8 8 E R C S D S D G L A 170 IBO $\begin{array}{c} 170 \\ 1150 \\ \text{CCT OCT CAG CAT CTT ATC CCA GTG GAA GGA AAT TTG COT GTG GTG GAG TAT TTG GAT GAC AGA AAC ACT TTT CCA CAT AGT GTG GTG GTG GCC P P Q E L I R V E G N L R V E Y L D D R H T P R E S V V V P 190 \\ 190 \\ 1250 \\ \end{array}$ TAT GAG CCG CCT GAG --- Intron 6 ~ 500 bp Y E P P E -GTT GGC TCT GAC TGT ACC ACC ATC CAC TAC AAC TAC ATG TGT AAC AGT TGC TGC ATG V G S D C T T I H Y N Y N C N S S C N 230 240 220 1 300 GOC GOC ATG AAC COG AGG COC ATC CTC ACC ATC ATC ATC ATC ACA CTG GAA GAC TOC AG --- Intron 7 ~350 bp G G N N R R P I L T I T L E D S S 250 260 1350 T GOT AAT CTA CTG GGA 250 A 1400 CGG AAC AOC TTT GAG CTG CGT GTT TOT OCC TGT CCT GGG AGA GAC CGG COC ACA GAG GAA GAG AAT CTC COC AAG AAA GOG GAG CCT CAC R H S P E V R V C A C P G R D R R T E E E H L R K K G E P H 270 H 280 1450 1450
CAC GAG CTG CCC CCA GGC ACC ACT AAG CGA G -- Intron 8 91 bp -- CA CTG CCC AAC AAC ACC ACC TCT CCC CAG CCA AAG AAG

B
C

B
C

B
C

B
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

C
C

< AAA CCA CTG GAT GGA GAA TAT TTC ACC CTT CAG--- Intron 9 ~2 500 bp --- ATC CGT GGG CGT GAG CGC TTC GAG ATG TTC CGA R P L D G E Y P T L Q I R G R E R P E N P R 330 340 E 1600 1650 TABOU GAG GOC TTG GAA CTC AAG GAT GOC CAG GOT GOG AAG GAG CCA GOG GOG AOC AGG CCT E A L E L K D A Q A G K E P G G S R A 350 360 - Intron 10 ~ 900 bp L H S 5 350 1700 CAC CTC AAG TCC AAA AAG GOT CAG TCT ACC TCC CGC CAT AAA AAA CTC ATG TTC AAG ACA GAA GOG CCT H L R S R R G Q S T S R H K R L M P K T E G P 370 380 1800 390 GAC TCA GAC TGA CATTCT 1800 380 1850 CETTOCTTOCAATAGGTOTOCOTCA 1950 TOCAT GAACAACTTGOCCTGCA TOGOGACTAGGACAT 2050 A GCATGTTTGGGAGATGTA CHARTER OCTAGOGGCCCACTTCACCG 2150 TTOCAC R.TGOO 2250 2300 CCTTGACCETTC CONTRACT. CAGTTGOG 2350 2400 COCCCARCCAAACCCTCT MATCTCACCCCATCCCACACCCTGGAGGATTTCATCTCTTGTATAT TGATCTCCATCC 2500 2550 ACCANGACTTGTTTTAT COT 2650 GECTTACTECAGE COCCOCTOC GETTCATO 2750 2700 CETCTCTCAGECTOCCAGAGTO GTCCAGCTGGAAGGGT 2850 ACCTUTUTUTUTUTUTUAGGOTG

FIG. 2. DNA sequence of the human p53 gene and the predicted amino acid sequence of the protein. Differences between this sequence and human p53 cDNA sequences are indicated above the relevant nucleotides and below the relevant amino acids. Nucleotides previously omitted from the sequence of the cDNA clone p102 (25) are overlined. The 5' and 3' ends of the cDNA p53J6K are indicated with arrowheads. The polyadenylation signal is underlined, and the nucleotide at which polyadenylation occurs (2929) is marked by a dot. Intron sizes preceded by a \sim are derived from restriction mapping data and are $\pm 10\%$; other intron lengths were obtained from sequence data and are exact. complete p53 gene, further genomic libraries were screened. A phage library, constructed by cloning partial HaeIII and AluI digests of human fetal liver DNA into λ Charon 4A (21) was screened, using HU2-6 and HU7-1 as probes. A cosmid library, which was prepared with MboI partial digest of DNA from the lymphoblastoid line MANN (38), was also screened with a probe derived from the 5' end of p53J6K. Two phage clones, $\lambda p53\alpha$ and $\lambda p53\pi$, and a single cosmid clone, cosp53PHD, were isolated and characterized by restriction mapping and Southern blotting to the cDNA clones. This showed that $\lambda p 53\alpha$ and $\lambda p 53\pi$ are overlapping clones covering 18 kb of DNA and that this region is, in turn, included within the 38-kb insert from cosp53PHD (Fig. 1). Comparison of genomic Southern blots with blots of phage and cosmid DNA indicated that no rearrangements occurred during cloning. Together with sequence analysis (see below), the Southern blot data indicate that cosp53PHD contains an intact p53 gene.

Organization and sequence of the human p53 gene. Restriction fragments containing exons were identified by hybridization to the cDNA clones p53J6K and p102, subcloned into M13 vectors, and sequenced. Comparison of the sequence so obtained with that of the cDNA clones showed that the gene is split into 11 exons which vary in size from 22 to 1,268 bp (Fig. 1). The sequence of the exons, the predicted amino acid sequence of the corresponding protein, and the sizes of the introns are given in Fig. 2. Intron lengths were obtained either from restriction mapping or from direct sequence analysis (Fig. 2). Overall, the organization of the human gene is similar to that of the functional mouse gene (6). However, the gene is much larger, covering at least 20 kb of DNA as opposed to 12 kb. This is due to the fact that some of the introns in the human gene are longer, most notably the large first intron, which is 10 kb compared with 6 kb, and intron 9, which is 2.5 versus 0.83 kb. The positions of the splice junctions have been conserved in all cases but one. The second exon in the human gene is 19 bp longer than that of the mouse gene due to an alteration in the splice acceptor site. In the human second exon, 28 bp of 5'-untranslated region precede the first in-frame methionine (Fig. 2), whereas in the mouse exon no 5'-untranslated region precedes the first in-frame methionine (6). As noted previously (17, 45), the first in-frame methionine codon in the human sequence is not the same as that in the mouse, because another ATG occurs further in the mouse sequence. If this served as the initiation codon the amino-terminal sequence of the mouse protein would be MTAMEE as opposed to MEE for the human protein. It is not known whether the first or second methionine is actually used to initiate translation of the mouse protein. In any case, the change in position of the splice acceptor site means that there are at least 19 nucleotides in the human 5'-untranslated region that have no equivalent in the mouse sequence. This discontinuity can be observed in published comparisons of the mouse and human cDNA sequences (45). The sequences of the splice donor and acceptor sites are all in good agreement with the proposed consensus sequences (29); in particular the invariant dinucleotides GT and AG found at the 5' and 3' ends of introns are conserved.

A comparison of the exon sequences obtained from the genomic clones with the sequence of the cDNA clones p53J6K and p102 (25) reveals two differences, both of which are in the 3'-untranslated region. An ACT triplet (nucleotides 2235 to 2237) and a G residue (nucleotide 2252) found in the genomic sequence (Fig. 2) are absent from the sequence of the cDNA p102. Re-examination of the sequence data for p102 revealed that these nucleotides are, in fact, present but were omitted from the published figure. A complete match therefore exists between the sequence of p102 and the genomic sequence.

Identification of a promoter adjacent to exon 1. We argued, by analogy with the functional murine p53 gene (see below), that exon 1 was the most 5' exon of the gene and tested sequences immediately upstream for promoter activity. Restriction fragments extending 5' from the XbaI site within exon 1 were cloned in front of the CAT gene, giving two constructs designated p53RXB and p53BXB (Fig. 3A). The XbaI site is known to be within the transcription unit because it is present in human p53 cDNA clones (17, 45), thus including it in these constructs ensures that the normal transcriptional initiation site(s) is present. Both constructs and two control plasmids, pSV2 (which contains the SV40 early promoter cloned in front of the CAT gene) and pSVOsma (which contains no promoter), were transfected into HeLa cells. The cells were harvested 48 h after transfection, and lysates were tested for CAT activity (14). The results of one such assay are shown in Fig. 3B. As expected, cells transfected with pSV2 contained abundant CAT activity, whereas those transfected with pSVOsma showed no activity (14). p53RXB and p53BXB are both able to promote easily detectable levels of CAT; thus, sequences sufficient to initiate transcription are present on the 350-bp BamHI to XbaI fragment. It seems unlikely that such a promoter would occur fortuitously, and we therefore conclude that this fragment contains the human p53 promoter. In repeated experiments, cells transfected with pSV2 always contained more CAT activity than those transfected with p53RXB or p53BXB. Therefore, the human p53 promoter appears to be weaker than the SV40 early promoter in these cells, although these experiments did not allow us to quantitate this difference.



FIG. 3. (A) Construction of plasmids containing human p53 sequences 5' of exon 1 linked to the CAT gene. A partial restriction map of the 3.8-kb *Eco*RI fragment from cosp53PHD is shown; the following enzymes are indicated: E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; X, *Xba*I. The *Hin*dIII site 3' of exon 1 is not shown. The 0.35-kb *Bam*HI to *Xba*I fragment and the 2.4-kb *Eco*RI to *Xba*I fragment were cloned into the *Sma*I site (denoted by an S) of pSVOsma, in the indicated orientation, to give p53BXB and p53RXB, respectively. (B) Assay of CAT activity in HeLa cells transfected with 20 µg of pSV2 (SV2), p53RXB (RXB), p53BXB (BXB) or pSVOsma (SVO). Abbreviations; CM, chloramphenicol; 1CM, 1-acetyl-chloramphenicol; 3CM, 3-acetyl-chloramphenicol.

Vol. 6, 1986

Primer extension of SV80 mRNA. To obtain an accurate size for human p53 mRNA and, therefore, the size of exon 1, primer extension experiments were performed. A synthetic oligonucleotide (complementary to nucleotides 638 to 659 in Fig. 2) labeled at its 5' end was annealed either to $poly(A)^+$ RNA from SV80 cells (39) or to tRNA and was extended using reverse transcriptase (Fig. 4). The extended products specific to the reaction with SV80 mRNA are indicated by arrows. The longest of these products is 346 ± 5 nucleotides, as measured by comparison with a coelectrophoresed M13 sequencing ladder (data not shown). Assuming that this product represents reverse transcription to the 5' end of the mRNA, it gives a predicted size of almost exactly 2.6 kb for human p53 mRNA, excluding the poly(A) tail. Furthermore, the size of exon 1 is predicted to be 218 bp, and the cap site is predicted to be at nucleotide 313 (Fig. 2), which is approximately in the middle of the BamHI to XbaI fragment found to be active in CAT assays (Fig. 3). This localization of the mRNA cap site is thus entirely consistent with the CAT assay data and supports the idea that exon 1 is, in fact, the first exon of the gene. The possible nature of the series of reverse transcriptase stops seen between the markers at 238 and 242 bp is discussed below.

DISCUSSION

Results of Southern blot analysis with a variety of probes indicated that there was a single p53 gene in the human genome, which has recently been mapped to chromosome 17 (3, 22). This is in contrast to the mouse genome, in which two p53 loci are present: the functional gene and an intron-

FIG. 4. Primer extension on human p53 mRNA. A synthetic oligonucleotide labeled at its 5' end was hybridized to SV80 mRNA (lane SV80) or yeast tRNA (lane tRNA) and extended with reverse transcriptase. The products were run on a polyacrylamide gel and autoradiographed. Lane M contains radioactive size markers (pBR322 cut with HpaII).

less pseudogene. We have presented here the cloning and detailed characterization of the human p53 gene. The gene is split into 11 exons that are distributed over 20 kb of DNA. Although it is considerably larger, the human gene is organized in a similar fashion to the functional mouse gene. There is, however, one difference in the splicing pattern in that the position of the acceptor site at the 5' boundary of exon 2 is altered. This change results in the presence of 19 nucleotides in the human 5'-untranslated region for which there is no equivalent in the mouse sequence. It is conceivable that this difference has some functional significance, for example, in the control of p53 mRNA translation. Alternatively, it may simply represent a neutral mutational event.

Recently, Wolf et al. (43) reported the isolation of human p53 cDNA and genomic clones, which they used to perform a heteroduplex analysis of the gene. In contrast to the results presented here, they concluded that the human gene had seven exons and was therefore organized differently from the mouse gene. However, they cloned only the two 16-kb EcoRI fragments and thus are missing exon 1, which is on the 3.8-kb *Eco*RI fragment. Also, heteroduplex analysis does not provide sufficient resolution to detect small introns and exons, and this has led to an underestimate of the number of exons present. A similar heteroduplex analysis of the functional mouse gene (47) had suggested the presence of only 8 exons, whereas sequence analysis showed that there were in fact 11 (6). It appears that the same exons and introns have been missed in the heteroduplex analysis of the human gene as the mouse gene.

The structure of the normal human p53 gene is of interest in that abnormal p53 expression in at least some tumors could be due to a genetic rearrangement, although a preliminary survey of human carcinomas failed to detect any gross rearrangements in the p53 gene (11). However, the promyelocytic leukemia cell line HL-60 has been shown to have a large deletion in the p53 gene (44). A comparison of the structure of the normal and truncated genes may reveal the mechanism by which the deletion occurred, as it did in the case of studies on a truncated low-density lipoprotein gene (23). Other examples of p53 gene inactivation have been observed in murine systems (30, 42); in none of these instances is it known why selection for p53 gene inactivation occurred.

The results of the CAT assays and the primer extension analysis are entirely consistent with the fact that the promoter and the 5' end of the gene are located within the 0.35-kb BamHI to XbaI fragment adjacent to exon 1. A comparison of the sequence of this 0.35-kb region with sequences from the 5' end of the functional mouse p53 gene (6) showed that a high degree of homology (70%) exists between them over most of this region (Fig. 5). It is almost this exact fragment from the mouse gene that has recently been shown to have activity in CAT assays (7), although suprisingly, a fragment extending 100 bp further in the 5' direction had no activity. This was interpreted as evidence for a negative regulatory element in the 5' region of the gene (7). The 2.4-kb EcoRI to XbaI fragment from the human gene, which includes the human equivalent of this element, however, is active in CAT assays. Either the human gene may be regulated differently than the mouse gene or the 2.4-kb fragment may contain additional positive regulators that are responsible for its function.

Attempts to map the 5' end of mouse p53 mRNA have proved inconclusive (6, 46), a fact attributed to the presence of stable secondary structure within the 5'-untranslated region of the mRNA. Analysis of the sequence of the same



GGATCCAGCTGAGAGCAAACGC	-AAAAGCTTTCTTCCTTCCACCCTTCATATTTGACACA
*** *** ***** *	
GGAAAGAGCACAGAGCTCAGAACAGTGGCGGTCCACTTACGA	TAAAAACTTAATTCTTTCCACTCTITATACTTGACACA
AT-GCAGGATTCCTCCAAAATGATTTCCACCAATTCTGCCCT	CACAGCTCTGGCTTGCAGAATTTTCCACCCCAAAATGT
****** ***** ** * ******** *** *****	******** * *** * * * * * * *
GAGGCAGGAGCCCTCCGAATCGGTTTCCACCCATTTTGCCCT	CACAGCTCTATATCTTAGACGACTTTTCA-CAAAGCGT
TAGTATCTACGGCACCAGGTCGGCGAGAATCCTGACTCTGCA	CCCTCCTCCCCAACTCCATTTCCTTTGCTTCCTCCGGC
TCCTGCTGAGGGCAACATCTCAGGGAGAATCCTGACTCTGCA	AGTCCCCGCCTCCATTTCTTGCCCTCAACC
▼	
AGGCGGATTACTTGCCCTTACTTGTCATGGCGACTGTCCAGC	TTTCTGCCAGGAGCCTCGCAGGGGTTGATGGGATTGGG
	TPTGTGCCAGGAGTCTCGCGGGGGTTGCTCGGATTCGG
GTTTTCCCCTCCCATGTGCTCAAGACTGGCGCTAAAAGTTTT	GAGCTTCTCAAAAGTCTAGA HUMAN
***********	****** * *
ACTTTCCCCTCCCACGTGCTCACC-CTGGCTAAAGTTCTG	TAGCTTCAGTTCATTGGGAC MOUSE

FIG. 5. Comparison of human and mouse p53 DNA sequences 5' of the first exon. The human sequence is that of the 0.35-kb BamHI to XbaI fragment (nucleotides 73 to 432 in Fig. 2). The mouse sequence is nucleotides -470 to -101 of Bienz et al. (6). Colons indicate identity between two nucleotides, and dashes indicate the points at which gaps were introduced to maximize the homology. Symbols: ∇ , position of the 5' end of the longest primer extension product obtained using human p53 mRNA; \triangle , position of the 5' end of the longest fragment protected from S1 nuclease by mouse p53 mRNA (6); \oplus , position of the major reverse transcriptase stop site and major S1 nuclease cut site at 24°C obtained with mouse p53 mRNA (6, 46). Overlined nucleotides correspond to the 5' ends of the shorter primer extension products obtained with human p53 mRNA (Fig. 4).

region of the human gene also revealed the potential to form stable secondary structures. In the primer extension experiment (Fig. 4) a series of reverse transcriptase stops that may reflect the presence of such a structure are seen. The positions of these reverse transcriptase stop sites on the sequence of the human gene are equivalent to the positions of the major reverse transcriptase stop site (46) and the major S1 nuclease cut site at 24°C (6) found near the 5' end of mouse p53 mRNA (Fig. 5). This provides evidence that human and mouse p53 mRNA contain similar structures in their 5'-untranslated regions and, therefore, that these structures may be important in the regulation of p53 expression.

It is provocative that the 5' end of the longest primer extension product obtained with human p53 mRNA and the 5' end of the longest fragment protected from S1 nuclease by mouse p53 mRNA (6) map to equivalent positions in their respective genes (Fig. 5). This concordance further suggests that this position represents the mRNA cap site in both human and mouse p53 genes and that the human p53 promoter is contained within the 230-bp region between the BamHI site and the presumptive cap site (Fig. 5). This region contains no correctly positioned TATA box, nor is it GC rich as are other promoters that lack TATA boxes (33, 40). It is possible, therefore, that the human and mouse p53 genes have a distinct class of promoter. However, further deletion and mRNA mapping experiments are required to define more precisely the elements that constitute the human p53 promoter before detailed comparisons can be made with other promoter types.

The availability of murine p53 cDNA and genomic clones has allowed an analysis of the phenotypic effects of abnormal p53 expression on a variety of cell types. This led to the findings that mouse p53, under certain circumstances, can cooperate with an activated *ras* gene in the transformation of primary cells (13, 32), extend the life-span of primary cells (18), and make established cells highly tumorigenic in nude mice (12). The isolation of an intact human gene and cDNA clones that span the protein-coding region now allows a study of the effects of transfecting the human p53 gene into different cell types, both under the control of its own and exogenous promoters.

ACKNOWLEDGMENTS

We are very grateful to H. Kataoka and M. Collins for providing the human cDNA library, S. Carson for the human cosmid library, and P. W. J. Rigby for the CAT vectors. We thank S. Benchimol, G. Matlashewski, and D. Pim for useful discussions; S. Darling, G. Matlashweski, L. Banks, and S. Tuck for comments on the manuscript; and C. Middlemiss for secretarial assistance.

LITERATURE CITED

- Auersperg, J. 1964. Long term cultivation of hypodiploid human tumor cells. J. Natl. Cancer Inst. 32:135–163.
- Benchimol, S., J. R. Jenkins, L. V. Crawford, K. Leppard, P. Lamb, N. M. Williamson, D. Pim, and E. Harlow. 1984. Molecular analysis of the gene for the p53 cellular tumor antigen, p. 383-391. In G. F. Vande Woude, A. J. Levine, W. Topp, and J. D. Watson (ed.), Cancer cells, vol. 2., Oncogenes and Viral Genes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Benchimol, S., P. Lamb, L. V. Crawford, D. Sheer, T. B. Shows, G. A. P. Bruns, and J. Peacock. 1985. Transformation associated p53 protein is encoded by a gene on human chromosome 17. Somat. Cell. Mol. Genet. 11:505-509.
- Benchimol, S., D. Pim, and L. V. Crawford. 1982. Radioimmunoassay of the cellular protein p53 in mouse and human cell lines. EMBO J. 1:1055-1062.
- 5. Benton, D. W., and R. W. Davis. 1975. Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–182.
- Bienz, B., R. Zakut-Houri, D. Givol, and M. Oren. 1984. Analysis of the gene encoding the murine cellular tumor antigen p53. EMBO J. 3:2179–2183.
- Bienz-Tadmor, B., R. Zakut-Houri, S. Libresco, D. Givol, and M. Oren. 1985. The 5' region of the p53 gene: evolutionary conservation and evidence for a negative regulatory element. EMBO J. 4:3209-3213.
- 8. Blin, N., and D. W. Stafford. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res 9:2303-2308.
- 9. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J.

Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18: 5294–5299.

- Crawford, L. V., D. C. Pim, E. G. Gurney, P. Goodfellow, and J. Taylor-Papadimitriou. 1981. Detection of a common feature in several human tumor cell lines—a 53,000 dalton protein. Proc. Natl. Acad. Sci. USA 78:41-45.
- 11. Crawford, L. V., D. Pim, and P. Lamb. 1984. The cellular protein p53 in human tumors. Mol. Biol. Med. 2:261-272.
- Eliyahu, D., D. Michalovitz, and M. Oren. 1985. Overproduction of p53 antigen makes established cells highly tumorigenic. Nature (London) 316:158–160.
- Eliyahu, D., A. Raz, P. Gruss, D. Givol, and M. Oren. 1984. Participation of p53 cellular tumor antigen in transformation of normal embryonic cells. Nature (London) 312:646–649.
- Gorman, C. M., L. F. Moffat, and B. M. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. 2:1044–1051.
- Graham, F. L., and A. J. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52:456-467.
- Hanahan, D., and M. Meselson. 1980. Plasmid screening at high colony density. Gene 10:63–67.
- Harlow, E., N. M. Williamson, R. Ralston, D. M. Helfman, and T. E. Adams. 1985. Molecular cloning and in vitro expression of a cDNA clone for human cellular antigen p53. Mol. Cell. Biol. 5:1601–1610.
- Jenkins, J. R., K. Rudge, and G. A. Currie. 1984. Cellular immortalisation by a cDNA clone encoding the transformation associated phosphoprotein p53. Nature (London) 312:651-654.
- 19. Land, H., L. F. Parada, and R. A. Weinberg. 1983. Tumorigenic conversion of embryo fibroblasts requires at least two cooperating oncogenes. Nature (London) 304:596-602.
- Lane, D. P., and L. V. Crawford. 1979. T antigen is bound to a host protein in SV40-transformed cells. Nature (London) 278:261-263.
- 21. Lawn, R. M., E. F. Fritsch, R. C. Parker, G. Blake, and T. Maniatis. 1978. The isolation and characterisation of linked δ and β -globin genes from a cloned library of human DNA. Cell 15:1157-1174.
- Le Beau, M. M., C. A. Westbrook, M. D. Diaz, J. D. Rowley, and M. Oren. 1985. Translocation of the p53 gene in t(15;17) in acute promyelocytic leukaemia. Nature (London) 316:826-828.
- Lehrman, M. A., W. J. Schneider, T. D. Sudhof, M. S. Brown, J. L. Goldstein, and D. W. Russell. 1985. Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. Science 227:140–146.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Matlashewski, G., P. Lamb, D. Pim, J. Peacock, L. Crawford, and S. Benchimol. 1984. Isolation and characterisation of a human p53 cDNA clone: expression of the human p53 gene. EMBO J. 3:3257-3262.
- Mercer, W. E., C. Avignolo, and R. Baserga. 1984. Role of the p53 protein in cell proliferation as studied by microinjection of monoclonal antibodies. Mol. Cell. Biol. 4:276-281.
- Mercer, W. E., D. Nelson, A. B. DeLeo, L. J. Old, and R. Baserga. 1982. Microinjection of monoclonal antibody to protein p53 inhibits serum-induced DNA synthesis in 3T3 cells. Proc. Natl. Acad. Sci. USA 79:6309–6312.
- Milner, J., and S. Milner. 1981. SV40-53K antigen: a possible role for 53K in normal cells. Virology 122:785–788.
- Mount, S. 1982. A catalogue of splice junction sequences. Nucleic Acids Res. 10:459–472.
- 30. Mowat, M., A. Cheng, N. Kimura, A. Bernstein, and S.

Benchimol. 1985. Rearrangements of the cellular p53 gene in erythroleukaemic cells transformed by Friend virus. Nature (London) **314**:633–636.

- Murray, N. E. 1983. Phage lambda and molecular cloning, p. 395-432. In R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Parada, L. F., H. Land, R. A. Weinberg, D. Wolf, and V. Rotter. 1984. Cooperation between gene encoding p53 tumor antigen and *ras* in cellular transformation. Nature (London) 312: 649-651.
- 33. Reynolds, G. A., S. K. Basu, T. F. Osbourne, D. J. Chin, G. Gil, M. S. Brown, J. K. Goldstein, and K. L. Luskey. 1984. HMG CoA reductase: a negatively regulated gene with unusual promoter and 5' untranslated regions. Cell 38:275-285.
- Ruley, H. E. 1983. Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. Nature (London) 304:602–606.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 36. Sarnow, P., Y. S. Ho, J. Williams, and A. J. Levine. 1982. Adenovirus E1b-58K tumor antigen and SV40 large tumor antigen are physically associated with the same 54Kd cellular protein in transformed cells. Cell 28:387–394.
- 37. Schneider, U., H. U. Schwenk, and G. Bornkamm. 1979. Characterisation of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma Int. J. Cancer 19:621-626.
- 38. Spielman, R. S., J. Lee, W. F. Bodmer, J. G. Bodmer, and J. Trowsdale. 1984. Six HLA-D region α-chain genes on human chromosome 6: polymorphisms and associations of DCα-related sequences with DR types. Proc. Natl. Acad. Sci. USA 81:3461-3465.
- 39. Todaro, G., H. Green, and M. R. Swift. 1966. Susceptibility of human diploid fibroblast strains to transformation by SV40 virus. Science 153:1252–1254.
- 40. Valerio, D., M. G. C. Duyvesteyn, B. M. M. Dekker, G. Weeda, T. M. Berkvens, L. van der Voorn, H. van Ormondt, and A. J. van der Eb. 1985. Adenosine deaminase: characterisation and expression of a gene with a remarkable promoter. EMBO J. 4:437-443.
- Weinberg, R. A. 1984. Cellular oncogenes. Trends. Biochem. Sci. 9:131-133.
- 42. Wolf, D., N. Harris, and V. Rotter. 1984. Reconstitution of p53 expression in a non-producer Ab-MuLV-transformed cell line by transfection of a functional p53 gene. Cell **38**:119–126.
- Wolf, D., Z. Laver-Rudich, and V. Rotter. 1985. In vitro expression of human p53 cDNA clones and characterisation of the cloned human p53 gene. Mol. Cell. Biol. 5:1887–1893.
- Wolf, D., and V. Rotter. 1985. Major deletions in the gene encoding the p53 tumor antigen cause lack of p53 expression in HL-60 cells. Proc. Natl. Acad. Sci. USA 82:790-794.
- Zakut-Houri, R., B, Bienz-Tadmor, D. Givol, and M. Oren. 1985. Human p53 cellular tumor antigen: cDNA sequence and expression in COS cells. EMBO J. 4:1251–1255.
- 46. Zakut-Houri, R., D. Givol, B. Bienz, A. Rogel, and M. Oren. 1984. Structure and expression of the genes encoding the murine p53 cellular tumor antigen, p. 393–401. In G. F. Vande Woude, A. J. Levine, W. Topp, and J. D. Watson (ed.), Cancer cells, vol. 2., Oncogenes and viral genes, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 47. Zakut-Houri, R. M. Oren, B. Bienz, V. Lavie, S. Hazum, and D. Givol. 1983. A single gene and a pseudogene for the cellular tumor antigen p53. Nature (London) 306:594–597.