An Ankyrin-like Protein with Transmembrane Domains Is Specifically Lost after Oncogenic Transformation of Human Fibroblasts*

Daniel Jaquemar, Thomas Schenker, and Beat Trueb‡
From the M. E. Müller Institute, University of Bern, CH-3010 Bern, Switzerland

We have identified a novel transformation-sensitive mRNA, which is present in cultured fibroblasts but is lacking in SV40 transformed cells as well as in many mesenchymal tumor cell lines. The corresponding gene is located on human chromosome 8 in band 8q13. The open reading frame of the mRNA encodes a protein of 1119 amino acids forming two distinct domains. The N-terminal domain consists of 18 repeats that are related to the cytoskeletal protein ankyrin. The C-terminal domain contains six putative transmembrane segments that resemble many ion channels. This overall structure is reminiscent of TRP-like proteins that function as store-operated calcium channels. The novel protein with an Mr of 130 kDa is expressed at a very low level in human fibroblasts and at a moderate level in liposarcoma cells. Overexpression in eukaryotic cells appears to interfere with normal growth, suggesting that it might play a direct or indirect role in signal transduction and growth control.

Fibroblasts transformed by tumor viruses are often used in the laboratory as a model system to investigate transformation-associated changes in the phenotype and in the adhesive properties of the cells (1, 2). Usually these fibroblasts exhibit a roundish shape and show reduced adhesion to their substratum similar to authentic tumor cells. In contrast to spontaneously occurring cancer cells, virally transformed fibroblasts have the advantage that the normal counterpart, the original, is available for comparative studies. A set of transformation-sensitive proteins that are expressed by normal fibroblasts but specifically repressed after oncogenic transformation has emerged from these studies. The best known example is fibronectin, an important adhesion protein of the extracellular matrix (3). It is likely that the down-regulation of fibronectin with the 5′ end of the cDNA, the RACE technique was applied using methyl mercuric hydroxide-denatured mRNA from human fibroblasts (IMR90) and the AmpliFINDER RACE kit (CLONTECH). Two synthetic oligonucleotide primers were prepared for this purpose corresponding to the cDNA sequence (reverse primer nucleotides 532–547, nested primer nucleotides 439–466). The sequences of the isolated cDNA clones were determined on both strands by the dyeoxy chain termination technique (11) using Sequenase 2.0 (U. S. Biological Corp.). To resolve ambiguities observed in the 5′ region, suitable restriction fragments were subcloned into the sequencing vectors M13mp18 and M13mp19 and subsequently sequenced in the presence of dITP instead of dGTP. All sequences were analyzed with the software computer package of the Genetics Computer Group (University of Wisconsin, WI) using default settings unless otherwise stated. The sequences were compared with all entries of the EBI data bank (release 55.0) and the Swissprot data bank (release 36.0).

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) Y10601.

* This study was supported by grants from the Swiss National Science Foundation (31–50571.97) and the Bernese Cancer League. 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.

† M. E. Müller Institute, University of Bern, P. O. Box 30, CH-3010 Bern, Switzerland. Fax: +41 31 632-4999; E-mail: trueb@mem.unibe.ch.

‡ M.E. Müller Institute, University of Bern, CH-3010 Bern, Switzerland. Fax: +41 31 632-4999; E-mail: trueb@mem.unibe.ch.

(Received for publication, October 23, 1998)

We have recently used this model system in combination with a subtractive cDNA cloning approach to identify novel transformation-sensitive proteins (4, 5). Our efforts led to the isolation of more than 40 cDNA clones that showed a dramatic reduction in their relative synthesis after oncogenic transformation. The proteins encoded by these clones could be grouped into four distinct classes (5): cytoskeletal proteins (e.g. vinculin, SM22), extracellular matrix proteins (e.g. fibronectin, collagen VI), proteolytic enzymes (e.g. collagenase, urokinase), and regulatory proteins (e.g. IGFBP-5, myosin kinase). Six novel gene products were discovered during our study, including a GTP-binding protein (4), a zyxin-like protein (6), and a novel serine protease (7).

Here we set out to characterize one of the new cDNA clones. This clone codes for an integral membrane protein with 18 ANK repeats. The polypeptide appears to combine the features of the cytoskeletal protein ankyrin with those of an ion channel. The function of the novel protein remains obscure, although there is circumstantial evidence that it might be involved in a growth regulatory pathway.

EXPERIMENTAL PROCEDURES

Molecular Cloning and Sequencing—A cDNA clone of 207 bp that originated from a subtracted cDNA library (4, 5) was labeled with [α-32P]dCTP by the random primed oligo-labeling method (8). This probe was used to screen approximately 0.5 × 109 recombinant phages of a commercial cDNA library, which had been prepared from human lung fibroblasts (HL1011, CLONTECH Laboratories) by the plaques hybridization technique (9). Positive colonies were picked, amplified, and subcloned into the plasmid pUC19 following standard procedures (10). To obtain the 5′ end of the cDNA, the RACE technique was applied using methyl mercuric hydroxide-denatured mRNA from human fibroblasts (IMR90) and the AmpliFINDER RACE kit (CLONTECH). Two synthetic oligonucleotide primers were prepared for this purpose corresponding to the cDNA sequence (reverse primer nucleotides 523–547, nested primer nucleotides 439–466). The sequences of the isolated cDNA clones were determined on both strands by the dyeoxy chain termination technique (11) using Sequenase 2.0 (U. S. Biological Corp.). To resolve ambiguities observed in the 5′ region, suitable restriction fragments were subcloned into the sequencing vectors M13mp18 and M13mp19 and subsequently sequenced in the presence of dITP instead of dGTP. All sequences were analyzed with the software computer package of the Genetics Computer Group (University of Wisconsin, WI) using default settings unless otherwise stated. The sequences were compared with all entries of the EBI data bank (release 55.0) and the Swissprot data bank (release 36.0).

Northern Blotting—Total RNA was extracted from confluent cell layers by the SDS/protease K method (12). For some experiments this RNA was further purified by chromatography on oligo(dT)-cellulose utilizing an mRNA isolation kit (Amersham Pharmacia Biotech). The RNA (15 μg total RNA or 2 μg of poly(A) RNA/lane) was resolved on 1% agarose gels in the presence of 1 M formaldehyde and transferred to nylon membranes by vacuum blotting (10). The membranes were hybridized with radiolabeled cDNA probes at 42 °C in a buffer containing 1 The abbreviations used are: bp, base pair(s); APH, aminoglycoside phosphotransferase; DAPI, 4,6-diamino-2-phenylindole; DMEM, Dulbecco’s modified Eagle’s medium; FISH, fluorescence in situ hybridization; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; TRP, transient receptor potential protein; TRPL, TRP-like protein; MOPS, 3-[N-morpholino]propanesulfonic acid.
50% formamide. After being washed at regular stringency (10), the membranes were exposed to BioMax MS film (Kodak, Rochester, NY) or analyzed with a phosphorimager (Storm 840, Molecular Dynamics, Sunnyvale, CA). The multiple tissue Northern blots used in this study contained poly(A) + RNA from 16 adult human tissues (MTN I, MTN II, CLON-TECH) and total RNA from 4 fetal human tissues (Northern Territory Human Fetal Tissue, Invitrogen, San Diego, CA). All these blots were processed as described above.

Quantitative PCR—Competitive PCR was performed essentially as described by Gilliland et al. (13). Poly(A)+ RNA (0.25 μg) was transcribed into cDNA with 25 units of Superscript II (Gibco-BRL) in the presence of 0.5 mU of RNaseH (Boehringer Mannheim). The reaction (total volume 20 μl) was primed with an oligonucleotide corresponding to positions 3115–3137 of the cDNA sequence. Aliquots of the single-stranded cDNA (0.5 μl) were amplified with AmpliTaq polymerase (Perkin Elmer) in the presence of 2.5 mM MgCl₂ through 35 cycles of 10 s at 94 °C, 10 s at 50 °C, and 10 s at 72 °C. The upper primer for this PCR corresponded to positions 2780–2802 of the cDNA, the lower primer was identical compared with the one used for first strand synthesis. For quantitative experiments, the PCR was competed with serial 3-fold dilutions of a competitor template. This competitor consisted of the SacI/PstI restriction fragment derived from the cDNA (positions 2682–3366) which contained, in its primer was identical compared with the one used for first strand synthesis (HTB 88), SK-UT-1 (HTB 114), and COS-1 (CRL-1650). All cells were (HTB 153), RD (CCL-136), HT1080 (CCL 121), Hs913T (HTB 152), WI38 VA13 (CCL 75.1), A204 (HTB 82), A673 (CRL 1598), Hs729T Collection (ATCC; Manassas, VA): WI38 (CCL75), IMR90 (CCL 186), human fibroblasts (4, 5). One of the clones derived from this library attracted our attention because its expression was completely down-regulated upon oncogenic transformation. This clone was 207-bp long and appeared to code for a novel ankyrin-like protein.

The clone was now used as a probe to screen a commercial cDNA library prepared from human fibroblasts with the intention to isolate the full coding region for the novel ankyrin-like protein. After four rounds of screening we obtained 21 independent A clones, but none of them reached all the way to the 5’ end of the corresponding mRNA. The RACE technique was therefore utilized to amplify the 5’ end, which resulted in the isolation of 11 additional clones.

Altogether, the 32 cDNA clones covered a contiguous sequence of 4230 bp (Fig. 1). This sequence started with a 5’ untranslated region of 174 bp that was particularly rich in the nucleotides G and C. The first ATG codon at position 175–177 was preceded by two in frame stop codons. This putative translation start site was followed by an open reading frame of 3357 nucleotides that terminated in the stop codon TAG (position 3532–3534). After the stop codon, there was a 3’ untranslated sequence of 708 nucleotides which harbored two consensus polyadenylation signals AATAAA at position 3842–3847 and 4218–4223. The sequence of the probe used for screening was found at position 1607–1813.

Amino Acid Sequence—The open reading frame predicts a protein of 1119 amino acids with a molecular mass of 127.4 kDa and an isoelectric point of 7.0 (Fig. 1). In the following, this polypeptide will be termed p120. It contains five potential glycosylation sites of the form NXT/S, but only two sites are likely to be used for modification (positions 747 and 753) because the amnylic acid sequence at the corresponding mRNA. The RACE technique was therefore utilized to amplify the 5’ end, which resulted in the isolation of 11 additional clones.

Altogether, the 32 cDNA clones covered a contiguous sequence of 4230 bp (Fig. 1). This sequence started with a 5’ untranslated region of 174 bp that was particularly rich in the nucleotides G and C. The first ATG codon at position 175–177 was preceded by two in frame stop codons. This putative translation start site was followed by an open reading frame of 3357 nucleotides that terminated in the stop codon TAG (position 3532–3534). After the stop codon, there was a 3’ untranslated sequence of 708 nucleotides which harbored two consensus polyadenylation signals AATAAA at position 3842–3847 and 4218–4223. The sequence of the probe used for screening was found at position 1607–1813.

Amino Acid Sequence—The open reading frame predicts a protein of 1119 amino acids with a molecular mass of 127.4 kDa and an isoelectric point of 7.0 (Fig. 1). In the following, this polypeptide will be termed p120. It contains five potential glycosylation sites of the form NXT/S, but only two sites are likely to be used for modification (positions 747 and 753) because the amnylic acid sequence at the corresponding mRNA. The RACE technique was therefore utilized to amplify the 5’ end, which resulted in the isolation of 11 additional clones.

Altogether, the 32 cDNA clones covered a contiguous sequence of 4230 bp (Fig. 1). This sequence started with a 5’ untranslated region of 174 bp that was particularly rich in the nucleotides G and C. The first ATG codon at position 175–177 was preceded by two in frame stop codons. This putative translation start site was followed by an open reading frame of 3357 nucleotides that terminated in the stop codon TAG (position 3532–3534). After the stop codon, there was a 3’ untranslated sequence of 708 nucleotides which harbored two consensus polyadenylation signals AATAAA at position 3842–3847 and 4218–4223. The sequence of the probe used for screening was found at position 1607–1813.

Amino Acid Sequence—The open reading frame predicts a protein of 1119 amino acids with a molecular mass of 127.4 kDa and an isoelectric point of 7.0 (Fig. 1). In the following, this polypeptide will be termed p120. It contains five potential glycosylation sites of the form NXT/S, but only two sites are likely to be used for modification (positions 747 and 753) because the amnylic acid sequence at the corresponding mRNA. The RACE technique was therefore utilized to amplify the 5’ end, which resulted in the isolation of 11 additional clones.

Altogether, the 32 cDNA clones covered a contiguous sequence of 4230 bp (Fig. 1). This sequence started with a 5’ untranslated region of 174 bp that was particularly rich in the nucleotides G and C. The first ATG codon at position 175–177 was preceded by two in frame stop codons. This putative translation start site was followed by an open reading frame of 3357 nucleotides that terminated in the stop codon TAG (position 3532–3534). After the stop codon, there was a 3’ untranslated sequence of 708 nucleotides which harbored two consensus polyadenylation signals AATAAA at position 3842–3847 and 4218–4223. The sequence of the probe used for screening was found at position 1607–1813.

Amino Acid Sequence—The open reading frame predicts a protein of 1119 amino acids with a molecular mass of 127.4 kDa and an isoelectric point of 7.0 (Fig. 1). In the following, this polypeptide will be termed p120. It contains five potential glycosylation sites of the form NXT/S, but only two sites are likely to be used for modification (positions 747 and 753) because the amnylic acid sequence at the corresponding mRNA. The RACE technique was therefore utilized to amplify the 5’ end, which resulted in the isolation of 11 additional clones.

Altogether, the 32 cDNA clones covered a contiguous sequence of 4230 bp (Fig. 1). This sequence started with a 5’ untranslated region of 174 bp that was particularly rich in the nucleotides G and C. The first ATG codon at position 175–177 was preceded by two in frame stop codons. This putative translation start site was followed by an open reading frame of 3357 nucleotides that terminated in the stop codon TAG (position 3532–3534). After the stop codon, there was a 3’ untranslated sequence of 708 nucleotides which harbored two consensus polyadenylation signals AATAAA at position 3842–3847 and 4218–4223. The sequence of the probe used for screening was found at position 1607–1813.

Amino Acid Sequence—The open reading frame predicts a protein of 1119 amino acids with a molecular mass of 127.4 kDa and an isoelectric point of 7.0 (Fig. 1). In the following, this polypeptide will be termed p120. It contains five potential glycosylation sites of the form NXT/S, but only two sites are likely to be used for modification (positions 747 and 753) because the amnylic acid sequence at the corresponding mRNA. The RACE technique was therefore utilized to amplify the 5’ end, which resulted in the isolation of 11 additional clones.

Altogether, the 32 cDNA clones covered a contiguous sequence of 4230 bp (Fig. 1). This sequence started with a 5’ untranslated region of 174 bp that was particularly rich in the nucleotides G and C. The first ATG codon at position 175–177 was preceded by two in frame stop codons. This putative translation start site was followed by an open reading frame of 3357 nucleotides that terminated in the stop codon TAG (position 3532–3534). After the stop codon, there was a 3’ untranslated sequence of 708 nucleotides which harbored two consensus polyadenylation signals AATAAA at position 3842–3847 and 4218–4223. The sequence of the probe used for screening was found at position 1607–1813.
27–55% similarity if conserved amino acid substitutions are included) with the consensus motif XGXXTPLHLAARXGHVEVVKLLLDXGADVNA

given this repetitive structure, it seems

FIG. 1. Nucleotide and derived amino acid sequence of clone p120. Cysteine residues are encircled, and potential glycosylation sites are marked by triangles. Open symbols indicate glycosylation sites located at the cytoplasmic side of the plasma membrane, and closed symbols those located at the extracellular side.
likely that the N-terminal portion of the novel protein evolved from an ancestral ANK motif that was duplicated 15–18 times during evolution. ANK motifs occur in a large variety of proteins with vastly different functions (21–23), including cytoskeletal proteins (ankyrin), transcriptional regulators (NFκB, IκB, bcl-3), and membrane-bound receptors (Notch).

Fig. 2. Topological model of p120. A, alignment of ANK repeats present in the p120 polypeptide. Identical residues are boxed. The repeats are ordered according to similarity (dendrogram). B, hydropathy plot of the p120 polypeptide calculated according to Kyte and Doolittle. Putative transmembrane domains are shaded. C, putative model of p120 in the cell membrane. ANK repeats are shown as shaded circles, and sequences sharing only partial similarity with the ANK consensus sequence are shown as open circles. Transmembrane domains are indicated by white bars. Asn-linked carbohydrates are denoted by “Y”.

### Table: ANK Sequence Alignment

| Consensus | 1 | 61 | 11 | 12 | 14 | 10 | 2 | 4 | 9 | 6 | 7 | 3 | 13 | 5 | 8 | 15 |
|-----------|---|----|----|----|----|----|---|---|---|----|---|---|----|---|---|----|
| ANK       | X | G | N | P | E | L | K | A | A | K | N | K | X | X | K | X | L | X | X | X |
| ANK 1     | D | M | D | F | T | P | L | B | A | A | K | E | G | C | O | L | H | E | K | T |
| ANK 2     | D | K | K | S | S | L | H | F | A | S | Y | G | R | I | N | T | C | O | R | L |
| ANK 3     | H | G | M | T | P | T | L | B | L | A | A | K | N | K | C | H | D | K | E | V | Q |
| ANK 4     | D | G | M | T | A | L | D | F | A | A | R | E | C | H | A | R | A | V | A | L | L |
| ANK 5     | D | G | N | T | A | L | D | F | A | A | R | E | C | H | A | R | A | V | A | L | S |
| ANK 6     | D | G | C | P | T | P | L | A | C | E | G | G | P | G | S | W | N | L | G | P | N | V | S | I | H | S | R | K |
| ANK 7     | Y | G | N | T | P | T | C | A | V | E | K | N | Q | I | S | V | K | F | L | S | G | R | G | A | N | P | N | L | R | N | F |
| ANK 8     | N | G | N | A | V | I | I | T | A | C | T | N | S | H | A | L | O | I | L | N | G | A | R | P | C | R | S | N | K |
| ANK 9     | N | G | R | S | P | T | L | B | A | T | A | S | A | S | W | H | I | V | L | L | S | G | K | A | Q | V | D | I | K | D |
| ANK 10    | G | K | A | P | P | B | L | A | D | Q | D | D | L | N | M | I | X | C | H | S | D | N | A | O | X | D | P | V | E | K | 270 |
| ANK 11    | G | K | A | P | P | B | L | A | D | Q | D | D | L | N | M | I | X | C | H | S | D | N | A | O | X | D | P | V | E | K | 270 |
| ANK 12    | G | K | A | P | P | B | L | A | D | Q | D | D | L | N | M | I | X | C | H | S | D | N | A | O | X | D | P | V | E | K | 270 |
| ANK 13    | G | K | A | P | P | B | L | A | D | Q | D | D | L | N | M | I | X | C | H | S | D | N | A | O | X | D | P | V | E | K | 270 |
| ANK 14    | G | K | A | P | P | B | L | A | D | Q | D | D | L | N | M | I | X | C | H | S | D | N | A | O | X | D | P | V | E | K | 270 |
| ANK 15    | G | K | A | P | P | B | L | A | D | Q | D | D | L | N | M | I | X | C | H | S | D | N | A | O | X | D | P | V | E | K | 270 |

Amino Acid Residues

Hydropathy

0 1 2 3

0 1 2 3
The C-terminal part of the novel polypeptide encompasses seven hydrophobic segments of about 20 amino acids each, which may act as transmembrane domains. A hydrophobicity plot according to Kyte and Doolittle (24) locates these segments at positions 722–742, 767–787, 806–824, 850–870, 873–893, 941–961, and 1010–1030 (Fig. 2B). Another program based on the algorithm of Engelman et al. (25) also predicts seven hydrophobic segments, but this program places one of them at position 897–917 instead of position 850–870. It is therefore possible that only six of the hydrophobic segments function as transmembrane domains, while one of them (probably the segment at position 850–870 which contains several charged residues) does not span across the entire membrane but enters the lipid bilayer partially. This structure is reminiscent of some ion channels, which possess six transmembrane domains and a pore loop gating the channel (26, 27). Since the polypeptide does not contain a typical signal peptide at its N terminus, an internal segment must serve as signal peptide as found for example with the anion exchanger of the red blood cell membrane (28, 29). The topological orientation of membrane proteins can be predicted by analysis of positive charges adjacent to the first transmembrane domain (30). These predictions suggest that the novel protein will assume a type II orientation, in which the N-terminal domain with all the ANK repeats lies at the cytoplasmic side of the cell membrane as proposed in Fig. 2C. In this model, the two glycosylation signals found at positions 747 and 753 occur in the first extracellular loop and should therefore be accessible for modification by N-linked carbohydrates. It is obvious that this proposed structure awaits verification by experiments.

In addition to the ANK repeats and to the transmembrane domains, the predicted protein contains an N-terminal segment of 60 residues and a C-terminal segment of 90 residues. Both of these segments do not reveal substantial similarity to any known protein motif.

**Similarity to Other Proteins and to Expressed Sequence Tags**—The amino acid sequence of the novel protein was compared with all entries of the Swissprot data bank. Numerous hits were found that corresponded to the cytoskeletal protein ankyrin and to other proteins containing ANK repeats. In addition to these hits, we found one protein that had emerged from the genome project of the nematode Caenorhabditis elegans (C29E6.2, accession number Z72504). Using a GAP creation weight of 9 and an extension weight of 3, an alignment with 28% identity or 37% similarity resulted and extended over the entire length of the two proteins, including the ANK repeats and the transmembrane domains (not shown). Nothing is known about the function of the *C. elegans* protein as it was simply predicted by computer analysis of the genomic sequence. Nevertheless, the striking conservation of the primary structure between two such distantly related species may indicate a fundamental function of the new protein.

The cDNA sequence of p120 was also compared with all expressed sequence tags stored in the EST data bank. Three partial cDNA clones were found that matched our cDNA sequence (not shown), namely AA502490 (441 bp), AA972567 (284 bp), and AA873172 (580 bp). Interestingly, all these clones were derived from human tumor samples (colon, lung, and kidney), which were analyzed in the course of the Cancer Genome Anatomy Project (CGAP). The significance of this finding is not yet clear.

**Expression of p120 in Cells and Tissues**—The expression of the novel gene was analyzed by Northern blotting experiments. On a blot containing RNA from different batches of human fibroblasts (WI38, IMR90), our cDNA clones hybridized to a single band of 4.6 kilobases (Fig. 3). This size is in agreement with the length of our cDNA sequence (4230 bp) assuming a poly(A) tail of about 300 nucleotides. With RNA from SV40 transformed fibroblasts (VA13) no signal was obtained, indicating that transcription of the corresponding gene was completely turned off after oncogenic transformation.

To determine whether expression of the p120 gene was repressed exclusively in virally transformed fibroblasts or whether this down-regulation was a common feature of transformed cells, we analyzed 11 different cell lines derived from spontaneous mesenchymal tumors (Fig. 3). Cells from four rhabdomyosarcomas, two fibrosarcomas, an osteosarcoma, and a chondrosarcoma did not possess any traces of the mRNA. When two different leiomyosarcoma cell lines were examined, the p120 mRNA was detected at a low level in one, but not in the other. Only cells from a liposarcoma contained large amounts of the p120 mRNA. Rehybridization of the same Northern blot with a probe for GAPDH demonstrated that all lanes contained similar amounts of intact RNA. Thus, the repression of the p120 mRNA is a common feature of most, but not all, mesenchymal tumor cells.

We now asked the question what tissues might express the novel mRNA. To this end our probes were hybridized to commercial Northern blots containing mRNA from various adult human tissues, including skeletal muscle, heart, lung, spleen, kidney, liver, pancreas, thymus, brain, prostate, testis, ovary, small intestine, colon, and cultivated leukocytes. To our surprise, we were not able to detect any signal that was stronger than background (not shown). Likewise, a Northern blot containing RNA from four fetal human tissues (brain, liver, lung, and skeletal muscle) did not yield any distinct signal. Rehybridization of the same blots with probes for β-actin or GAPDH produced strong signals, indicating that all the blots contained normal amounts of undegraded mRNA. It therefore appears that all the human tissues investigated above produce extremely low amounts of the p120 mRNA, which cannot be traced by Northern blotting.

We therefore used a very sensitive PCR approach to detect expression of the p120 gene in normal human tissues (Fig. 4).
The mRNA was transcribed into cDNA and used as template for competitive PCR. A pair of strand-specific primers was selected, which allowed amplification of a 358-bp fragment corresponding to the fourth and fifth putative transmembrane domain. These domains were chosen because we have evidence that they are encoded by two separate exons. For quantitation, a competitor template was used as an internal standard. This template consisted of the same 358-bp sequence but harbored additionally an unrelated sequence of 52 bp (total length 410 bp) in its center to permit distinction between competitor DNA and the 358-bp cDNA on a polyacrylamide gel. The competitor template was added to the reactions at serial 3-fold dilutions. At a high cDNA to competitor ratio, amplification of the 358-bp fragment was observed, whereas at a low ratio, amplification of the 410-bp competitor occurred. The absolute amount of cDNA in the reaction mixture could thus be estimated from the transition point where competitor DNA and cDNA fragment were obtained in equimolar amounts. Applying this technique to mRNA from cultivated fibroblasts, the expected band of 358 bp was obtained (Fig. 4). Equimolar amounts of the 358-bp fragment and the 410-bp competitor were reached when 3.3 x 10^6 copies of the competitor template were added to the reaction mixture. This number would correspond to approximately 500 copies of p120 mRNA per fibroblast. When RNA from a 12-week old human embryo was analyzed in a similar way, the diagnostic band of 358 bp was also obtained, indicating that the p120 gene was in fact expressed in human tissues. Competition experiments suggested that these tissues contained at least a 1000-fold lower level of p120 mRNA than cultivated fibroblasts. RNA from embryonic muscle possessed an even lower level, which was barely detectable by this sensitive method (not shown). The same results were obtained when another set of PCR primers was used. Thus, the p120 gene is faithfully transcribed in human tissues but at an extremely low level.

Localization of the p120 Gene—To further characterize the p120 gene, we set out to determine its localization in the human genome by the FISH technique. When a biotinylated probe prepared from our cDNA clones was hybridized to metaphase spreads of human chromosomes, a clear signal was observed on the long arm of chromosome 8 (Fig. 5). Among 100 mitotic spreads examined, 89 showed specific signals on at least one pair of chromatids. The exact position as determined by superimposing the FISH signal with the DAPI banding pattern was found to be band 8q13. No additional locus was observed. The gene for the novel protein is therefore situated at a single locus on human chromosome 8. It might be of interest to note that the gene for ankyrin 1 is also located on this chromosome, in region 8p11.2 (23).

Initial Characterization of the Protein—In an effort to characterize the protein encoded by the p120 gene, polyclonal antibodies were raised against a synthetic peptide comprising 15 amino acids from the C terminus of the predicted polypeptide (residues 1105–1119). In enzyme-linked immunosorbant assays, the antibodies recognized their antigen even at high dilution (not shown). On the other hand, they did not work for radioimmunoprecipitations or indirect immuno-fluorescence studies probably because they were directed against a sequential epitope, which may not be accessible in the folded protein. When tested on Western blots, the antibodies reacted readily with a bacterially expressed fusion protein spanning the intracellular, C-terminal domain of p120 (Fig. 6). We therefore used Western blotting to analyze protein extracts from liposarcoma cells, which possess a relatively high level of the p120 mRNA as demonstrated above. A total cell extract did not yield any specific immunoreactive band, but an extract enriched for membrane proteins showed a major band with an apparent molecular mass of 130 kDa and several minor bands with molecular masses smaller than 60 kDa (Fig. 6). It is likely that the 130-kDa band corresponds to the full-length product of the p120 mRNA, whereas the minor bands may represent unrelated proteins cross-reacting with our antibodies as they were obtained also with all other samples examined. When a membrane extract from fibroblasts was analyzed in a similar way, a very faint band of 130 kDa was detected (not shown). In contrast, no bands in the 130-kDa region were observed with membrane extracts from SV40 transformed fibroblasts or HT1080 fibrosarcoma cells. Thus, liposarcoma cells and fibroblasts produce p120 at a low but clearly detectable level.

Overexpression in Eukaryotic Cells—For a functional analysis of the novel protein, we tried to express large amounts of p120 in human cells. To this end, a full-length cDNA construct was transfected into A204 and HT1080 cells. Positive transfectants were selected by their resistance to the antibiotic G418, and resistant colonies were tested on Northern blots for their expression of the p120 mRNA (Fig. 7). Only a very few resistant colonies were obtained, and more surprisingly, only one of them (n = 32) was found to express a mRNA related to p120. However, a closer inspection revealed that even this colony did not express the full-length mRNA, but rather a truncated form of less than 3000 nucleotides that cannot encode a functional protein (Fig. 7, lane 9). Rehybridization of the same Northern blot with a cDNA probe for the G418 resistance gene AHPAP demonstrated that all resistant cells did in fact express the AHPAP mRNA. When a batch of freshly transfected cells was investigated prior to subcloning, a weak signal corresponding to the p120 mRNA was initially observed. However, this signal gradually disappeared after subcultivation, whereas the AHPAP signal persisted. Similar results were obtained when transient expression studies were performed with COS-1 cells. We therefore concluded, that expression of the p120 mRNA is selectively turned off in transfected cells, either by specific removal of the p120 cDNA or by selective inactivation of the incorporated gene. One possibility to explain these results is that the uncontrolled expression of p120 is not compatible with the normal growth of human cells.
DISCUSSION

By subtractive cDNA cloning, we have identified a novel gene on human chromosome 8 that codes for a transformation-sensitive protein termed p120. Although we have utilized state-of-the-art immunological and molecular biological techniques, we have not yet been able to demonstrate the exact function of this protein. The biochemical characterization was made difficult by the unusual regulation of the novel gene product as follows.

The mRNA for p120 was expressed by cultured fibroblasts, but it was specifically repressed after oncogenic transformation. Based on a quantitative PCR analysis, we estimate that a fibroblast contains about 500 copies of the mRNA, which corresponds to a gene that is expressed at moderate level. In contrast, a variety of mesenchymal tumor cell lines did not express p120 at all, with the exception of liposarcoma cells that contained a relatively high level of the mRNA.

The protein encoded by the p120 mRNA could barely be detected in fibroblasts, but it was found at a moderate level in liposarcoma cells. Since these cells possess a fairly high level of

Fig. 5. Chromosomal localization of the p120 gene. Metaphase spreads of human lymphocytes were hybridized with a labeled cDNA probe for p120 (A). The chromosomes were counterstained with DAPI (B). The gene was assigned to chromosomal band 8q13 by superimposing the FISH signal and the DAPI banding pattern (C).

Fig. 6. Detection of p120 by antibodies. A crude preparation of a fusion protein containing amino acid residues 1063–1119 (lanes 1 and 4), as well as two cell membrane extracts, one obtained from liposarcoma cells (SW872, lanes 2 and 5), the other from SV40 transformed fibroblasts (VA13, lanes 3 and 6) were resolved on a 3–10% gradient SDS-polyacrylamide gel and either stained with Coomassie Blue (lanes 1–3) or transferred to nitrocellulose and stained with polyclonal antibodies (lanes 4–6). The antibodies recognize a sequential epitope of p120 within amino acids 1105–1119. The migration positions of globular protein standards are shown in the left margin.

Fig. 7. Transfection experiments. A full-length construct for p120 was transfected into A204 cells. The entire population of freshly transfected cells (lane 3) as well as individual colonies selected by their resistance to G418 (lanes 4–9) were analyzed by Northern blotting. The blot was hybridized with a cDNA probe for p120 (top) or a cDNA probe for the APH resistance gene (bottom). The migration positions of the ribosomal RNA subunits are indicated in the right margin. RNA from IMR90 fibroblasts (lane 1) and from A204 cells (lane 2) was included as control.
the mRNA, the protein must either turn over very quickly or it is translated with very poor efficiency. There is evidence in favor of the latter possibility. Analysis of the mRNA reveals an extremely high proportion of rare amino acid codons along the entire sequence.

In contrast to cultivated fibroblasts, p120 was expressed at a very low level in human tissues. These levels could not be detected by Northern blotting, but required a very sensitive PCR approach for detection. Based on a competitive PCR experiment, we estimate that the tissues of a human embryo possess a 1000-fold lower level of the mRNA than cultivated fibroblasts. It is possible that the gene is transcribed only at a very restricted area by specialized cells. However, our Northern blotting experiments with 20 adult and embryonic tissues and our preliminary studies with in situ hybridization do not support this idea. It seems more likely that most cells express the gene at a very low level and/or for only a very short time during their life span.

The low expression of p120 is reflected by the fact that the comprehensive data bank of expressed sequence tags contains only three partial clones that overlap with our sequence. All these clones are derived from human tumor samples. Overexpression of the p120 cDNA in normal and transformed cells has not been successful. Even utilization of an ecdysone-inducible system did not overcome the problem. A formed cells has not been successful. Even utilization of an these clones are derived from human tumor samples. All the structural features of p120 are now shared by members of a loosely defined family of proteins, the TRP-like molecules (for review, see Refs. 31 and 32). Originally, TRP and its related protein TRPL were discovered in the fruit fly Drosophila as G-protein-regulated channels that mediate the light-activated conductance in the visual system. Like p120, TRPL is composed of about 1100 amino acids, it lacks an N-terminal signal peptide, but contains six trans-membrane domains and a pore loop, and it possesses four ANK repeats in its N-terminal domain. Recent evidence suggests that TRP and TRPL form part of a capacitative (or store-operated) channel (31, 32). Such channels are activated by depletion of Ca2+ from internal stores. They are widely distributed in invertebrates and vertebrates and seem to be located in the plasma membrane of most excitable and nonexcitable cells. Several TRP-related proteins have also been identified in humans (e.g. TRPC1 and TRPC3). All these TRP-like molecules have a topology similar to p120, although they show a rather low similarity at the amino acid level. TRPC3 (htrp3, Ref. 33) taken as example shares 27% identity or 39% similarity with p120 over a region of 364 amino acids corresponding to part of the ANK repeats and part of the transmembrane domains. This similarity may suggest that p120 represents a novel member in the heterogeneous family of TRP-like molecules.

To conclusively demonstrate that p120 is indeed an ion channel, it is inevitable to express our novel protein in a eukaryotic expression system. So far, all our efforts in this direction have been fruitless since the uncontrolled expression of p120 did not seem to be compatible with normal growth of the transfected cells. However, it is conceivable that a selected domain (rather than the entire protein) could successfully be expressed in eukaryotic cells. Thus, experiments performed with suitable fragments derived from our cDNA clones might eventually shed light on the function and regulation of this interesting protein.

Acknowledgment—We thank Dr. K. H. Winterhalter for continuous support.

REFERENCES

1. Pastan, I., and Willingham, M. (1978) Nature 274, 645–650
2. Dubbelman, T. M. A. R., and Yamada, K. M. (1982) Biochim. Biophys. Acta 693, 177–187
3. Hynes, R. O. (1990) Fibronectins, Springer-Verlag, New York
4. Schenker, T., Lach, C., Kessler, B., Calderara, S., and Trube, B. (1994) J. Biol. Chem. 269, 25447–25453
5. Schenker, T., and Trube, B. (1998) Exp. Cell. Res. 239, 161–168
6. Zumbrunn, J., and Trube, B. (1998) Eur. J. Biochem. 241, 657–663
7. Zumbrunn, J., and Trube, B. (1998) FEBS Lett. 439, 187–192
8. Feinberg, A., and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13
9. Benton, W. D., and Davis, R. W. (1977) Science 196, 180–182
10. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, S. J., and Struhl, K. (1987) Current Protocols in Molecular Biology, Greene
An Ankyrin-like Protein with Transmembrane Domains

11. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
12. Rowe, D. W., Moen, R. C., Davidson, J. M., Byers, P. H., Bornstein, P., and Palmieri, R. D. (1978) Biochemistry 17, 1581–1590
13. Gilliland, G., Perrin, S., Blanchard, K., and Bunn, H. F. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2725–2729
14. Baltensperger, K., and Porzig, H. (1997) J. Biol. Chem. 272, 10151–10159
15. Heng, H. H. Q., Squire, J., and Tsui, L.-C. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9509–9513
16. Zumbrunn, J., and Trueb, B. (1997) Genomics 45, 461–462
17. Trueb, B., and Winterhalter, K. H. (1996) EMBO J. 5, 2815–2819
18. Harlow, E., and Lane, D. (1988) Antibodies, a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
19. Laemmli, U. K. (1970) Nature 227, 680–685
20. Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4350–4354
21. Bennett, V. (1992) J. Biol. Chem. 267, 8703–8706
22. Michaely, P., and Bennett, V. (1992) Trends Cell Biol. 2, 127–129
23. Peters, L. L., and Lux, S. E. (1993) Semin. Hematol. 30, 85–118
24. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–152
25. Engelman, D. M., Steitz, T. A., and Goldman, A. (1986) Annu. Rev. Biophys. Chem. 15, 321–353
26. MacKinnon, R. (1995) Neuron 14, 889–892
27. Jan, L. Y., and Jan, Y. N. (1992) Cell 69, 715–718
28. Tanner, M. J. A. (1993) Semin. Hematol. 30, 34–57
29. Tam, L. Y., Loo, T. W., Clarke, D. M., and Reithmeier, R. A. F. (1994) J. Biol. Chem. 269, 32542–32550
30. Sipos, L., and von Heijne, G. (1993) Eur. J. Biochem. 213, 1333–1340
31. Berridge, M. J. (1995) Biochem. J. 312, 1–11
32. Montell, C. (1997) Mol. Pharmacol. 52, 755–763
33. Zhu, X., Jiang, M., Peyton, M., Boulay, G., Hurst, R., Stefani, E., and Birnbaumer, L. (1996) Cell 85, 661–671