Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 5178–5182, July 1989
Neurobiology

Molecular cloning and primary structure of human glial fibrillary acidic protein

(cDNA cloning/astrocytes/intermediate filaments/cytoskeleton)

STEVEN A. REEVES, LEE J. HELMAN, AUDREY ALLISON, AND MARK A. ISRAEL

Molecular Genetics Section, Pediatric Branch, National Cancer Institute, Bethesda, MD 20892

Communicated by D. Carleton Gajdusek, February 17, 1989 (received for review August 1, 1988)

ABSTRACT **Glial fibrillary acidic protein (GFAP) is an intermediate-filament (IF) protein that is highly specific for cells of astroglial lineage, although its tissue-specific role is speculative. Determination of the primary structure of this protein should be of importance for understanding the functional role it plays in astroglia. Therefore, we isolated a cDNA clone encoding this protein and determined its nucleotide sequence. The predicted amino acid sequence indicates that GFAP shares structural similarities—particularly in the central rod domain and to a lesser degree in the carboxyl-terminal domain—with other IF proteins found in nonepithelial cell types. Considerable sequence divergence in the amino-terminal region of GFAP suggests that the tissue-specific functions of this IF protein might be mediated through this region of the molecule. In contrast, conservation of structural characteristics and a moderate degree of sequence conservation in the carboxyl-terminal region suggest functional similarities. Blot hybridization analysis using the GFAP cDNA as a probe failed to detect GFAP mRNA in both normal and neoplastic human tissues in which IF proteins other than GFAP are known to be expressed.**

Intermediate-filament (IF) proteins are a family of proteins that share common structural features and contribute to the formation of the cytoskeleton (for recent reviews, see refs. 1–3). These filaments consist of insoluble, fibrous polypeptides 8–11 nm in diameter, a size that is between that of microtubules and actin filaments, the other known families of cytoskeletal proteins. Among the IF proteins, many different keratins have been identified in epithelial tissues, but only a few different IF proteins have been identified in nonepithelial tissues: neurofilament proteins in neurons, desmin in muscle tissues, glial fibrillary acidic protein (GFAP) in astrocytes, and vimentin in tissues arising from embryonic mesenchyme. The monomeric protein subunits of IFs in a variety of tissues from different species have been evaluated, and the recognition of several highly conserved features has led to a consensus structural model for IF proteins (4, 5).

In every case examined to date, it has been found that IF proteins contain three distinct domains (1–3). The most prominent of these is a highly conserved central "rod" domain of ≈310 amino acids. Although the actual DNA sequence identity between the regions of IF genes encoding rod domains may range from 30% to 70% (5), the structural organization of this IF protein segment is invariant. The rod domain contains two similarly sized regions separated by a short spacer loop of ≈30 amino acids (5). Each of these two regions is characterized by a repeated pattern of 7 amino acid residues that are capable of forming α-helical arrays (2). In sharp contrast to the highly conserved central rod domains of IF proteins, the amino- and carboxyl-terminal regions of the different IF proteins have distinctive amino acid sequence content and widely divergent structural features (1–3, 6).

Although two different IF proteins may rarely be coexpressed in a single cell type, in most cases cells of any individual lineage express only one specific IF protein (7). Recognition of this cell-type-specific regulation of IF protein gene expression has led to the hypothesis that the tissue-specific features of IF protein function will ultimately be found to reside in the widely diverged, variable terminal regions of the molecule (2). Although GFAP is a specific marker of cells in the astrocytic lineage, virtually nothing is known regarding possible unique functions of this particular IF protein. An understanding of the primary structure of this IF protein will be crucial for our understanding of such functions, and we have therefore determined the amino acid sequence of human GFAP (hGFAP) predicted from a cDNA clone we have isolated.*

In addition to a large body of biochemical data, several previous reports have provided data relevant for an understanding of the primary structure of hGFAP. Geisler and Weber (8) have sequenced ≈50% of the porcine GFAP (pGFAP) protein, and partial cDNA and genomic molecular clones of murine GFAP (mGFAP) have been evaluated (9, 10). Also, partial DNA sequence data have been reported for hGFAP, corresponding to regions coding for the rod domain (11, 12). As anticipated, our data document considerable homology between GFAPs from different species. The predicted amino-terminal sequence of hGFAP, a region in which the tissue-specific functions of this protein might be determined, agrees closely with the protein sequence that has been determined for the porcine protein, although it varies considerably from that predicted for the putative murine protein.

MATERIALS AND METHODS

Cell Culture. Tumor cell lines were grown in Dulbecco's modified Eagle's medium supplemented with either 10% fetal bovine serum (GIBCO) or 10% newborn calf serum (GIBCO), 2 mM L-glutamine, and antibiotics (100 units of penicillin and 50 μg of streptomycin per ml) at 37°C in 5% CO_2. The human glioma cell lines used were HTB17 (U-373 MG) (13), U-138 MG (13), U-343 MGA CL 2:6 (13), LN18 (14), LN215 (14), U-251 MG-O (13), and HTB138 (Hs 683) (13).

Construction of cDNA Library and DNA Sequencing. Total cellular RNA was isolated from the cell line HTB17 by the guanidinium thiocyanate/cesium chloride procedure (15) and poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography (16). Double-stranded cDNA was prepared (17) from 5 μg of poly(A)$^+$ RNA. cDNA molecules were inserted into λgt10 DNA (18) and packaged *in vitro*, and duplicate

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

Abbreviations: IF, intermediate filament; GFAP, glial fibrillary acidic protein; hGFAP, human GFAP; mGFAP, murine GFAP; pGFAP, porcine GFAP.

*The sequence reported in this paper is being deposited in the GenBank data base (accession no. J04569).

filters of a portion of the library (10^5 plaques) were then screened with a ^{32}P-labeled mGFAP DNA probe (9). Recombinant phage DNA was prepared (19) and purified DNA fragments were subcloned into pGEM-3Z (Promega) by standard procedures (20). The cDNA clone phgp-1 was sequenced directly from plasmid DNA (21) by the chain-termination method (22) with Sequenase (United States Biochemical) as the DNA polymerase. Sequencing was initiated on both strands by using SP6 and T7 promoter primers (Promega). Subsequently, the entire phgp-1 insert was sequenced on both strands by using synthetic oligonucleotides as primers.

Isolation of DNA and Southern Blot Analysis. DNA was extracted by standard procedures (20) from tissue and cell-line pellets that had been stored at −70°C. For Southern blot analysis, DNA (15 μg) was digested with *Eco*RI restriction endonuclease (Bethesda Research Laboratories) according to the supplier's recommendations, electrophoresed in 0.8% agarose gels, and transferred to Nytran filters (Schleicher & Schuell). Hybridizations were carried out essentially as described by Southern (23). The mGFAP cDNA *Eco*RI insert (9) and phgp-1 *Eco*RI insert probes were labeled *in vitro* to a specific activity of ≈2.5 × 10^8 cpm/μg by nick-translation using [α-^{32}P]dCTP. The final hybridization wash was with 15 mM NaCl/1.5 mM sodium citrate, pH 7/1% NaDodSO$_4$ at 65°C for 1 hr.

Isolation of RNA and Northern Blot Analysis. Total cellular RNA was prepared (15) from human tissues and cell lines. RNA (20 μg per lane) was size-fractionated by electrophoresis in 1% agarose/6% formaldehyde gels, transferred to Nytran, hybridized, and washed as described (24).

RESULTS

To identify molecular clones encoding hGFAP, we constructed a cDNA library in λgt10 by using poly(A)$^+$ RNA from the human glioma-derived cell line HTB17, which is known to express high levels of GFAP (25). Using a cDNA probe that recognizes DNA sequences encoding mGFAP (9), we screened ≈10^5 λgt10 recombinants and obtained 5 phage clones hybridizing to the probe. All 5 recombinants were then subcloned into pGEM-3Z and found to have restriction endonuclease patterns compatible with the possibility that they contained overlapping stretches of DNA. One of these subclones, phgp-1, contained an insert of ≈3 kilobases (kb), a length similar to that of the GFAP transcript we observed in RNA from HTB17 by Northern blot hybridization (data not shown) using a probe that recognizes mGFAP mRNA (9). The complete nucleotide sequence of this clone was determined for both DNA strands and is shown in Fig. 1, where it is aligned with the DNA sequence of human vimentin (26), the only other human type III IF protein (3) for which the complete nucleotide sequence is known. Fig. 2 shows a

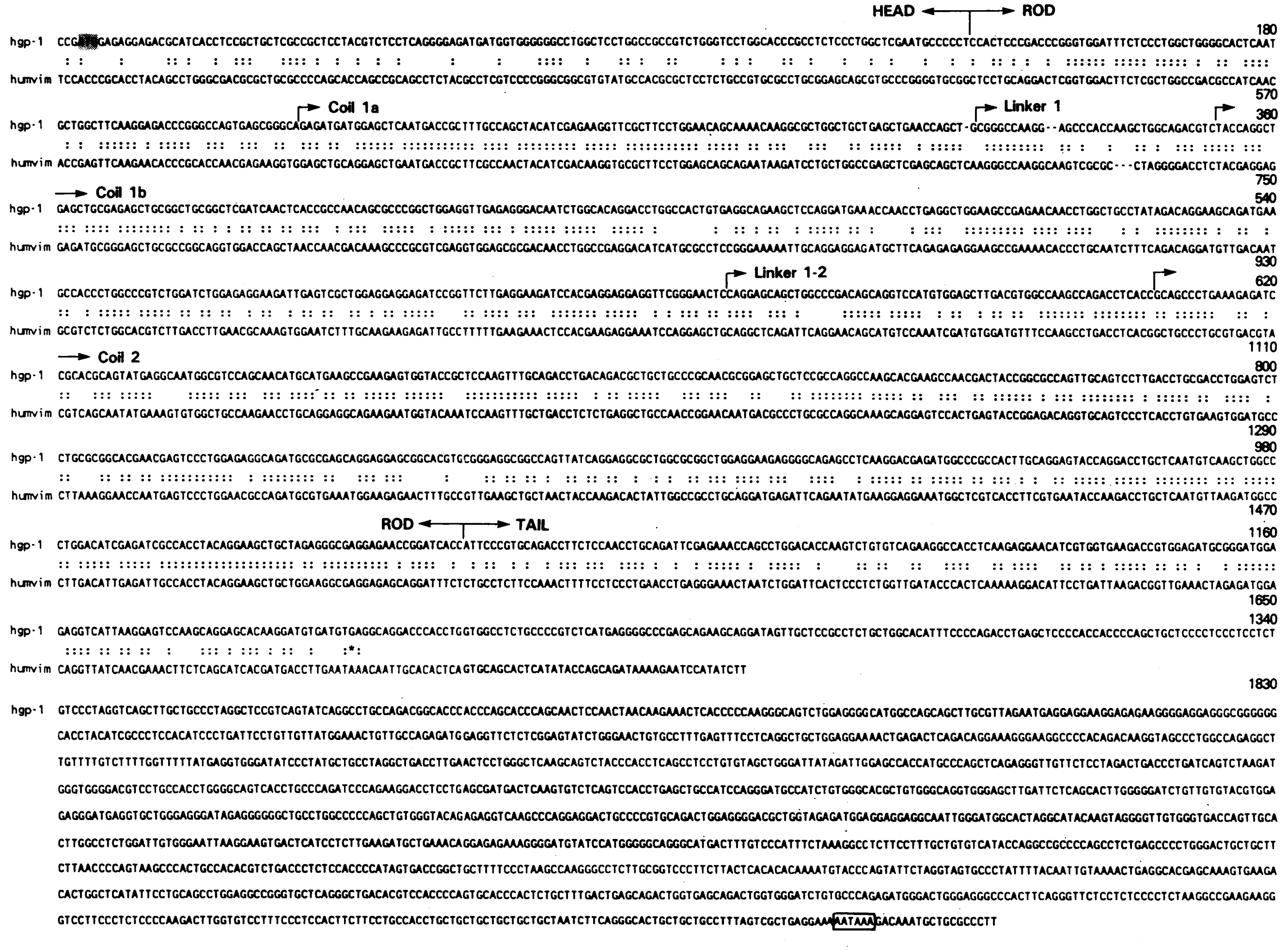

FIG. 1. Comparison of the nucleotide sequences of hGFAP (hgp-1) and human vimentin (humvim) (26). Nucleotides at the end of each line are numbered. In the GFAP sequence, dashes represent gaps inserted to optimize the alignment. A shaded box indicates the putative initiation methionine codon of GFAP, and an open box indicates a possible polyadenylylation signal. Regions coding for the predicted protein domains in the mature proteins are indicated [head, rod (consisting of coil 1a, linker 1, coil 1b, linker 1-2, and coil 2), and tail] above the alignment. Identical nucleotides between the two sequences are signified by colons and the translation termination codon is indicated by an asterisk.

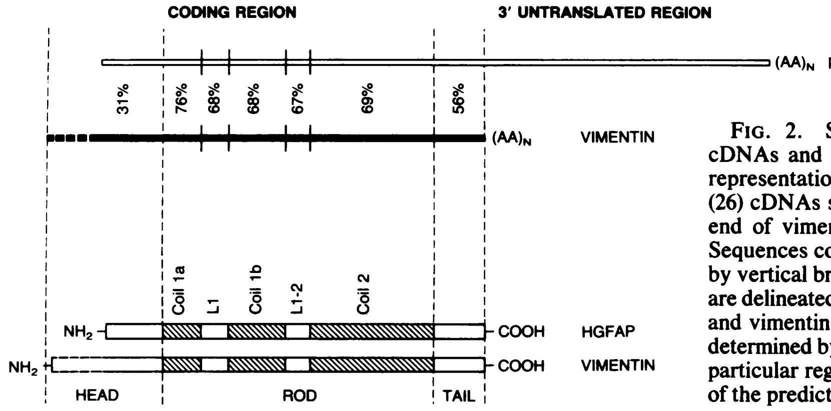

FIG. 2. Sequence homology between phgp-1 and vimentin cDNAs and between their predicted proteins. (*Upper*) Schematic representation of the alignment between phgp-1 and human vimentin (26) cDNAs shown in Fig. 1. The broken black bar shown at the 5' end of vimentin cDNA indicates sequence not shown in Fig. 1. Sequences coding for the head, rod, and tail domains are delineated by vertical broken lines, and those coding for subdomains of the rod are delineated by vertical solid lines. The homologies between phgp-1 and vimentin coding sequences are shown as percentages and were determined by the ratio of identical bases to total number of bases for particular regions of the cDNAs. (*Lower*) Schematic representation of the predicted hGFAP and vimentin. L1, linker 1; L1-2, linker 1-2.

schematic representation of this alignment and summarizes the nucleotide identities found between human vimentin and GFAP. A comparison of the nucleotide sequences of phgp-1 and the coding regions of the human vimentin gene indicates that the region of greatest identity lies in those sequences encoding the rod domain (Fig. 2 *Upper*). The overall identity of sequences in this region of vimentin cDNA and the corresponding region of phgp-1 is 66%.

A comparison at the nucleic acid level of the amino- and carboxyl-terminal regions of human vimentin with the corresponding regions of phgp-1 reveals the expected sequence divergence. There is also rather modest structural variation in the head region, where the human vimentin sequence extends an additional 99 base pairs (bp) upstream from the putative initiation codon for hGFAP (data not shown, but see Fig. 2 and ref. 26). In the tail region, the genes for these proteins share an important structural feature; namely, both have translation termination codons located at identical positions, compatible with the assertion that all IF genes arose from a common progenitor (2, 27, 28).

There is an extensive 3' untranslated region in phgp-1 (1734 bp), while the human vimentin gene has a relatively short 3' untranslated region of 57 bp. Within this region the hGFAP cDNA contains a putative polyadenylylation signal located 18 bp upstream from the poly(A) tract, whereas human vimentin has two possible polyadenylylation signals, one within the 3' untranslated region, located 20 bp upstream from the poly(A) tract, and one that overlaps the translation termination codon. It has not been determined which of these polyadenylylation signals is actually used.

An alignment of the amino acid sequences of GFAP from different species is shown in Fig. 3. The primary structure of pGFAP was determined by amino acid sequence analysis of GFAP-derived peptides and corresponds to ≈50% of the porcine protein (8). The amino acid sequences for hGFAP and mGFAP were predicted from our human cDNA clone, phgp-1, or from mouse cDNA and genomic DNA sequences reported by others (9, 10). The homology between the predicted amino acid sequence of hGFAP and the porcine amino-terminal head domain and carboxyl-terminal tail do-

```
                                           head<-||->rod
pgfap   MERRRVTSAARRSYVSSLVTVGGG----RRLGPGPRLSLARMPPPLPARVDFSLAGALNT  56
hgfap   MERRRITSAARRSYVSSGEMMVGGLAPGRRLGPGTRLSLARMPPPLPTRVDFSLAGALNA  60
mgfap                  MPPRRWSGASGPSRQLGTMPRFSLSRMPTPLPARVDFSLAGALNT  45

                  ->coil 1a                           ->l1      ------->
pgfap   GFKETRASERAEMMELNDRFASYIEL                                    85
hgfap   GFKETRASERAEMMELNDRFASYIEKVRFLEQQNKALAAELNQLRAKEPTKLADVYQAEL 120
mgfap   GFKETRASERAEMMELNDRFASYIELVRFLEQQNKALAAELNQLRAKEPTKLADVYQAEL 105

        ------------->coil 1b
hgfap   RELRLRLDQLTANSARLEVERDNLAQDLATVRQKLQDETNLRLEAENNLAAYRQEADEAT 180
mgfap   RELRLRLDQLTANSARLEVERDNLAQLAGTLRQKLQDETNLRLEAENNLAAYRQEAHEAT 165

                                   ->l1-2                 ------->
hgfap   LARLDLERKIESLEEEIRFLRKIHEEEVRELQEQLARQQVHVELDVAKPDLTAALKEIRT 240
mgfap   LARVDLERKVESLEEEIQFLRKIYEEEVRDLREQLAQQQVHVEMDVAKPDLTAALREIRT 225

        ------------->coil 2
pgfap                                                        CEVDSLR
hgfap   QYEAMASSNMHEAEEWYRSKFADLTDAAARNAELLRQAKHEANDYRRQLQSLTCDLESLR 300
mgfap   QYEAVATSNMQETEEWYRSKFADLTDAASRNAELLRQAKHEANDYRRQLQALTCDLESLR 285

pgfap   GTNESLERQMREQEERHAREAASYQEALTRLEEEGQSLKDEMARHLQEYQELLNVKLALD
hgfap   GTNESLERQMREQEERHVREAASYQEALARLEEEGQSLKDEMARHLQEYQDLLNVKLALD 360
mgfap   GTNESLERQMREQEERHARESASYQEALTRLEEEGQSLKEEMARHLQEYQDLLNVKLALD 345

                   rod<-||->tail
pgfap   IEIATYRKLLEGEENRITIPVQTFSNLQIRETSLDTKSVSEGHLKRN--VKTVEMRDGEV
hgfap   IEIATYRKLLEGEENRITIPVQTFSNLQIRETSLDTKSVSEGHLKRNIVVKTVEMRDGEV 420
mgfap   IEIATYRKLLEGEENRITIPVQTFSNLQIRETSLDTKSVSEGHLKRNIVVKTVEMRDGEV 405

pgfap   IKESKQEHKDV
hgfap   IKESKQEHKD-VM                                                432
mgfap   IKDSKQEHKDVVM                                                418
```

FIG. 3. Amino acid sequence of the hGFAP (predicted from the hgp-1 insert) and its homology with porcine and murine GFAP. mGFAP is the protein predicted from mouse genomic sequences (10). The pGFAP sequence was determined by peptide sequence analysis and represents ≈50% of the pig protein (8). Amino acid residues are numbered on the right. Domains in the mature protein are indicated and regions of homology are indicated by shading. Dashes represent gaps introduced to optimize the alignment. l1, linker 1; l1-2, linker 1-2.

main was 83% and 99%, respectively, demonstrating the anticipated close relationship between these proteins and providing strong evidence that phgp-1 encodes hGFAP. Other experiments, in which a monoclonal antibody recognizing authentic hGFAP (Boehringer Mannheim) was used to immunoprecipitate protein translated *in vitro* from the single RNA species transcribed *in vitro* from phgp-1, complement these findings (data not shown).

Our observation that the pGFAP initiation methionine agrees with the putative initiation methionine designated for phgp-1 strongly suggests that phgp-1 encodes the entire hGFAP. We therefore predict that phgp-1 encodes a protein composed of 432 residues with a calculated molecular weight of 49,891. While the predicted amino acid sequence of hGFAP and the determined amino acid sequence of pGFAP are very similar in both the head and tail domains, there are major amino acid sequence differences between the head domains of GFAP from these species and that predicted for mGFAP (Fig. 3). The rod and tail domains of all three species are >90% homologous, indicating that this apparent divergence is limited to the amino-terminal region of the protein. The putative initiation methionine codon in the mGFAP gene (9) is located downstream from the initiation methionine codon that was determined for pGFAP and that is apparently conserved in hGFAP. Comparison of phgp-1 and mGFAP at the nucleic acid level indicates that this initiation methionine codon is also conserved in the mouse (10), although in a different reading frame, suggesting that the amino-terminal domains of GFAP from each of these species are more closely related than a cursory analysis would reveal. Alternatively, it is possible that the GFAP of these species may vary or that there are multiple, structurally divergent forms of GFAP either encoded by different genes or transcribed from the same gene but into different mRNAs. Since the pattern of hybridization of phgp-1 to human genomic DNA (Fig. 6, see below) suggests that there is only a single human gene for GFAP, the last possibility appears more likely.

We used phgp-1 DNA to examine the steady-state levels of GFAP mRNA in a variety of human tissues and tumor-derived cell lines. Fig. 4 demonstrates findings representative of this analysis and confirms a large body of immunohistochemical data indicating that GFAP expression is restricted to normal and neoplastic tissues of the central nervous system (29, 30). Our finding that GFAP expression may be regulated at the level of gene transcription extends these studies and suggests that an understanding of the molecular mechanisms that regulate the tissue-specific expression of GFAP might provide insight into the regulated expression of other genes unique to the astrocytic lineage.

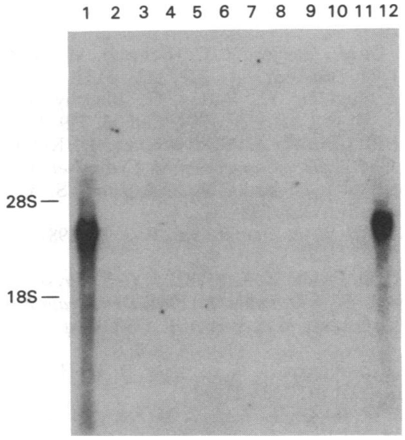

FIG. 4. Northern blot of total RNA from normal and malignant human tissues probed with [^{32}P]phgp-1. Lanes: 1, adult brain white matter; 2, kidney; 3, spleen; 4, liver; 5, muscle; 6, fibroblast line; 7, Ewing sarcoma cell line; 8, T-lymphoma cell line; 9, neuroblastoma cell line; 10 and 11, non-small-cell and small-cell lung carcinoma cell lines; 12, glioblastoma cell line HTB17. Positions of 28S and 18S rRNAs are shown.

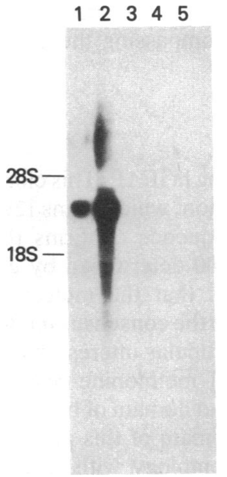

FIG. 5. Northern blot analysis of GFAP expression in glioma-derived cell lines. Total RNA (20 μg per lane) from the cell lines HTB17 (lane 1), U-343 MGA CL 2:6 (lane 2), HTB138 (lane 3), LN18 (lane 4), and LN215 (lane 5) was probed with [^{32}P]phgp-1.

The immunohistochemical detection of GFAP is one of the few objective means by which nonneuronal tumors of neuroepithelial origin can be classified and graded (7). Virtually all tumors thought to arise from astroglia have been found to stain positively for GFAP when examined by immunohistochemical techniques. Among the numerous tumor cell lines that have been established from GFAP-positive central nervous system tumors, however, we have examined 32 and found only 3 glioma cell lines, HTB17, U-343 MGA CL 2:6, and U-251 MG-O, that stain positively for this IF protein (data not shown). To determine whether an RNA encoding GFAP could be detected in cell lines we knew to be immunohistochemically negative for GFAP, we used phgp-1 as a probe to analyze RNA isolated from a number of such human glioma-derived cell lines. Fig. 5 shows a representative experiment from this survey in which we were unable to detect GFAP mRNA in three cell lines that were immunohistochemically negative for GFAP (lanes 3–5), whereas GFAP mRNA was easily detected in two cell lines known to stain positively for GFAP (lanes 1 and 2). These findings suggest that GFAP expression in glioma cell lines may be regulated at the level of transcription.

Gene rearrangements play an important role in the pathogenesis of human tumors (31, 32). For both lymphoid and myeloid tumors of the hematopoietic lineage, tumor-specific chromosomal rearrangements occur within genetic loci containing genes that are typically expressed at high levels in that particular cellular lineage (33). To determine whether the chromosomal region encoding GFAP might be rearranged during the development of glial tumors or tumor cell lines and thereby contribute to the lack of GFAP expression in these tissues, we analyzed the structural organization of the GFAP locus in several human glioma cell lines by Southern blot hybridization using [^{32}P]phgp-1 as a probe (Fig. 6). We observed two *Eco*RI fragments that were indistinguishable in the normal and neoplastic tissues. These findings indicate

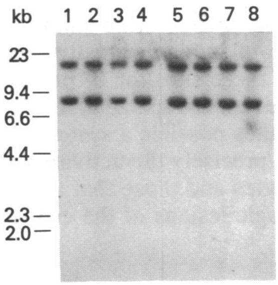

FIG. 6. Southern blot analysis of the hGFAP gene in normal tissues and glioma-derived cell lines. DNA samples (15 μg) were digested with *Eco*RI and probed with [^{32}P]phgp-1. *Hin*dIII-digested λ phage DNA provided size markers. Lanes: 1, adult brain; 2, fetal brain; 3, kidney; 4, HTB17; 5, U-343 MGA CL 2:6; 6, HTB138; 7, LN18; 8, LN215.

that no large genomic rearrangements have occurred within the region recognized by a probe encompassing the entire GFAP coding region.

DISCUSSION

We have isolated a cDNA clone encoding hGFAP. This clone contains the complete GFAP coding region, which spans 1296 nucleotides. The predicted protein sequence confirms the approximate molecular weight of 50,000 determined by gel electrophoresis (34, 35) and indicates that the molecular organization of GFAP is consistent with the consensus model proposed for IF proteins (2, 4). Of particular interest is our determination that the amino-terminal methionine residue found in pGFAP is conserved in the head domain of hGFAP. This finding indicates that the head domain of this protein, consisting of 45 residues, has little homology with, and is considerably shorter than, the comparable regions of either vimentin or desmin, which are the IF proteins from nonepithelial tissues that share the greatest homology with GFAP (2, 4). This variation is in contrast to the tail domains, in which there is both extensive primary sequence homology and structural similarity among these proteins.

To date, there has been little information indicating how IF proteins are assembled into filaments, although the basic structural unit of nonepithelial filaments seems to be a homodimer of IF protein (1, 2). Since isolated rods do not form authentic filaments (2), it has been suggested that the terminal, nonhelical domains of the molecule may play an important role in filament formation (2). Rarely, GFAP and vimentin are present in the same cell (36, 37). If copolymers between these two IF proteins can be formed, as suggested by *in vitro* studies (38, 39), the very disparate sizes of the terminal regions of GFAP and vimentin suggest it is more likely that the individual IF protein molecules would be arranged head-to-head and tail-to-tail rather than in a head-to-tail manner. This molecular organization would be consistent with the antiparallel arrangement of IF dimers proposed on the basis of structural analysis of α-keratin (40) and could provide both homopolymers and heteropolymers not only the capacity for increased structural variability but also the potential for enhanced functional plasticity. That vimentin and GFAP are coexpressed in immature glial cells (36), some glioma cell lines (38, 39), and reactive astrocytes (37), but not in mature, growth-arrested astrocytes, is compatible with this possibility and suggests that in each of these cell types the cytoskeleton might have quite distinct biological functions.

Another structural feature of the predicted hGFAP sequence that may be of functional importance is the conservation of serine and threonine residues in the carboxyl-terminal tail. These amino acids are potential sites for IF phosphorylation, a posttranslational modification known to occur in GFAP (41). IF protein phosphorylation has been most extensively studied for the neurofilament protein NF-H (42, 43), in which altered phosphorylation of the tail region is associated with alterations in the biophysical properties of the molecule (44). The conservation of five serine and threonine residues in the tail regions of human vimentin and GFAP suggests the importance of these sites for IF protein function, although their precise role remains speculative. The availability of a hGFAP cDNA makes possible a systematic genetic approach to defining more precisely the activities of this protein in both normal astrocytes and those that are so frequently associated with pathologic lesions of the central nervous system.

This work was supported in part by a grant from the Preuss Foundation.

1. Lazarides, E. (1982) *Annu. Rev. Biochem.* **51**, 219–250.
2. Weber, K. & Geisler, N. (1985) *Ann. N.Y. Acad. Sci.* **455**, 126–143.
3. Steinert, P. M., Steven, A. C. & Roop, D. R. (1985) *Cell* **42**, 411–419.
4. Geisler, N., Kaufmann, E. & Weber, K. (1982) *Cell* **30**, 277–286.
5. Osborn, M. & Weber, K. (1986) *Trends Biochem. Sci.* **11**, 469–471.
6. Lazarides, E. (1980) *Nature (London)* **283**, 249–256.
7. Osborn, M. (1983) *J. Invest. Dermatol.* **81**, 1045–1075.
8. Geisler, N. & Weber, K. (1983) *EMBO J.* **2**, 2059–2064.
9. Lewis, S. A., Balcarek, J. M., Krek, V., Shelanski, M. & Cowan, N. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2743–2746.
10. Balcarek, J. & Cowan, N. J. (1985) *Nucleic Acids Res.* **13**, 5527–5543.
11. Salim, M., Rehman, S., Sajdel-Sulkowska, E. M., Chou, W.-G., Majocha, R. E., Marotta, C. A. & Zain, S. B. (1988) *Neurobiol. Aging* **9**, 163–171.
12. Rataboul, P., Faucon Biguet, N., Vernier, P., De Vitry, F., Boularand, S., Privat, A. & Mallet, J. (1988) *J. Neurosci. Res.* **20**, 165–175.
13. Bigner, D. D., Bigner, S. H., Ponten, J., Westermark, B., MaHaley, M. S., Ruoslahti, E., Herschman, H., Eng, L. F. & Wikstrand, C. J. (1981) *J. Neuropathol. Exp. Neurol.* **40**, 201–226.
14. Tribolet, N. (1985) *Acta Neuropathol.* **66**, 208–217.
15. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
16. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
17. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
18. Grundmann, U., Amann, E., Zettlmeissl, G. & Kupper, H. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8024–8028.
19. Young, R. A. & Davis, R. W. (1983) *Science* **222**, 778–782.
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
21. Chen, E. Y. & Seeger, P. H. (1985) *DNA* **4**, 165–170.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
24. Helman, L. J., Thiele, C. J., Linehan, W. M., Nelkin, B. D., Baylin, S. B. & Israel, M. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2336–2339.
25. Rutka, J. T., Giblin, J. R., Dougherty, D. Y., Liu, H. S., McCulloch, J. R., Bell, C. W., Stern, R. S., Wilson, C. B. & Rosenblum, M. L. (1987) *Acta Neuropathol.* **75**, 92–103.
26. Ferrari, S., Battini, R., Kaczmarek, L., Rittling, S., Calabretta, B., de Riel, J. K., Philiponis, V., Wei, J.-F. & Baserga, R. (1986) *Mol. Cell. Biochem.* **6**, 3614–3620.
27. Quax, W., Egberts, W. V., Hendriks, W., Quax-Jensen, Y. & Bloemendal, H. (1983) *Cell* **34**, 215–223.
28. Marchuk, D., McCrohon, S. & Fuchs, E. (1984) *Cell* **39**, 491–498.
29. Bock, E. (1978) *J. Neurochem.* **30**, 7–14.
30. Bignami, A., Dahl, D. & Rueger, D. C. (1980) *Adv. Cell. Neurobiol.* **1**, 285–310.
31. Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. (1984) *Science* **224**, 1121–1125.
32. Taub, R., Moulding, C., Battey, G., Murphy, W., Vasicek, T., Lenoir, G. M. & Leder, P. (1984) *Cell* **36**, 339–345.
33. Kirsch, I. R., Brown, J. A., Lawrence, J., Korsmeyer, S. J. & Morton, C. C. (1985) *Cancer Genet. Cytogenet.* **18**, 159–171.
34. DeArmond, S. J., Fajardo, M., Naughton, S. A. & Eng, L. F. (1983) *Brain Res.* **262**, 275–282.
35. Schlaepler, W. W. & Zimmerman, V.-J. P. (1981) *Neurochemistry* **6**, 243–255.
36. Yen, S. H. & Fields, K. L. (1981) *J. Cell. Physiol.* **88**, 115–126.
37. Pixley, S. K. R. & DeVellis, J. (1984) *Dev. Brain Res.* **15**, 201–210.
38. Scharp, G., Osborn, M. & Weber, K. (1982) *Exp. Cell Res.* **141**, 385–395.
39. Wang, E., Carincross, J. G. & Liem, R. K. H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2102–2106.
40. Crewther, W. G., Inglis, A. S. & McKern, N. M. (1978) *Biochem. J.* **173**, 365–371.
41. Browning, E. T. (1985) *Trans. Am. Soc. Neurochem.* **16**, 21.
42. Julien, J. P. & Mushynski, W. E. (1983) *J. Biol. Chem.* **258**, 4019–4025.
43. Carden, M. J., Schlaepfer, W. W. & Lee, V. M. Y. (1985) *J. Biol. Chem.* **260**, 9805–9817.
44. Julien, J. P. & Mushynski, W. E. (1982) *J. Biol. Chem.* **257**, 10467–10470.