Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 5769–5773, July 1992
Plant Biology

Gene expression in cotton (*Gossypium hirsutum* L.) fiber: Cloning of the mRNAs

(tissue specificity/DNA sequence/promoter/particle bombardment/transgenic plants)

MALIYAKAL E. JOHN* AND LAURA J. CROW

Agracetus, Inc., 8520 University Green, Middleton, WI 53562

Communicated by Winston J. Brill, March 23, 1992

ABSTRACT **Cotton, an important natural fiber, is a differentiated epidermal cell. The number of genes that are active in fiber cells is similar to those in leaf, ovule, or root tissues. Through differential screening of a fiber cDNA library, we isolated five cDNA clones that are preferentially expressed in fiber. One of the cDNA clones, pCKE6, corresponded to an abundant mRNA in fiber. Transcripts for E6 were detected throughout the development of the fiber. Immunoprecipitation of *in vitro* translation products and Western blot analysis of fiber proteins showed two polypeptides in the range of 30–32 kDa as the products of E6 mRNA. Sequence analysis and hybrid-selected RNA translation also suggest that E6 mRNAs encode two polypeptides. Concentrations of E6 mRNA and protein are highest during the late primary cell wall and early secondary cell wall synthesis stages. Sequence comparison of E6 with other known eukaryotic and prokaryotic genes reveals no significant homology (GenBank; December 1991). E6 or a homologous gene(s) is conserved in several members of *Malvaceae* as well as in one other fiber-producing plant, kapok, but is not found in several other plants examined or in *Acetobacter xylinum*. A genomic clone corresponding to pCKE6 was isolated, and the promoter element of the E6 gene was shown to direct the expression of a carrot extensin mRNA in a tissue-specific and developmentally regulated fashion in transgenic cotton plants.**

Cotton is the most important textile fiber crop; therefore, the development of cotton fiber and the biosynthesis of cellulose are significant areas of research. Cotton fiber is a differentiated single epidermal cell of the outer integument of the ovule. It has four distinct growth phases; initiation, elongation (primary cell wall synthesis), secondary cell wall synthesis, and maturation (for a review, see ref. 1). Initiation of fiber development appears to be triggered by hormones, auxins, and gibberellins (for a review, see ref. 2). The primary cell wall is laid down during the elongation phase, lasting up to 25 days postanthesis (DPA). Synthesis of the secondary wall commences prior to the cessation of the elongation phase and continues to ≈40 DPA, forming a wall (5–10 μm) of almost pure cellulose. Upon maturity, cotton fiber is ≈87% cellulose. Thus cell wall biosynthesis is a major synthetic activity in fiber cells. The cell wall components are synthesized and transported by a functionally integrated membrane system of endoplasmic reticulum, Golgi complex, and plasmalemma. Newly synthesized cell wall polypeptides are released into the endoplasmic reticulum lumen before transportation and incorporation into cell walls. Glycosylation of structural proteins, as well as polymerization of hemicellulose and pectin, takes place in the Golgi complex from which the products are released into plasmalemma through the fusion of vesicles (1, 3).

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.

Plants, as well as certain bacteria, synthesize cellulose through enzyme systems localized in plasma membranes that use UDP-glucose as a substrate (1). The enzymatic pathway of cellulose synthesis in *Acetobacter xylinum* has been studied extensively and an operon encoding cellulose synthase has been cloned and sequenced (4, 5). In *A. xylinum* the cellulose is an extracellular secretion, whereas in plants it is a component of the cell wall. In addition to cellulose, cotton fiber also contains smaller amounts of pectins, hemicellulose, waxes, proteins, and inorganic salts. The mechanism by which cellulose microfibrils are produced and assembled along with the other components is not fully understood. Microtubules that are situated in the cytoplasm directly adjacent to the developing cell wall may participate in microfibril organization (6). Some of the protein constituents of cotton fiber (enzymatic, structural, or regulatory), as shown in this report, are unique to fiber cells and are likely to influence the development and properties of cotton. Evaluation of proteins from fibers of different developmental stages and other cotton plant tissues by two-dimensional gel electrophoresis has revealed fiber-specific proteins that are developmentally regulated (7). In the present study, we have identified and characterized a protein that is predominantly expressed in fiber cells.† Fiber-specific genes and their regulatory elements will be valuable tools to understand fiber differentiation and properties. In addition, this information is useful in genetic engineering for cotton fiber improvement.

MATERIALS AND METHODS

Identification and Characterization of Fiber-Specific Clones. Cotton plants (*Gossypium hirsutum* L. cv. Coker 312) were grown in the greenhouse. For fiber collection, flowers were tagged on the day of anthesis and bolls were collected at various stages of fiber development. Fibers removed from the ovules, young leaves, roots, ovules, and flowers (petal, ovule, stamen, and pistil) were stored frozen until use. The frozen materials were homogenized with a Polytron at full speed for 1.5 min in homogenization buffer added at a ratio of 1:4 (wt/vol). The homogenization buffer contained 5 M guanidine isothiocyanate, 0.2 M Tris acetate (pH 8.5), 0.7% 2-mercaptoethanol, 1% polyvinylpyrrolidone, and 0.62% sodium lauroyl sarcosine. The homogenate was subjected to cesium chloride centrifugation. DNA was prepared by the method of Dellaporta *et al.* (8). The extraction buffer was modified to contain 1% polyvinylpyrrolidone. Single-stranded cDNA was synthesized using oligo(dT) primers and Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories) as described by D'Alessio *et al.* (9). Homopolymer dC tails were added to double-stranded cDNA, which was then annealed with dG-tailed pBR322

Abbreviation: DPA, days postanthesis.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M92051).

(Bethesda Research Laboratories) for transformation of *Escherichia coli* RR1.

The cDNA library (1.5×10^4 clones) was subjected to differential screening with ^{32}P-labeled single-stranded cDNA probes from poly(A)$^+$ RNAs of leaf, ovule, flower, and 15- and 24-DPA fiber and root. Inserts of cDNA clones were subcloned into phagemid vector Bluescript II SK$^+$ (Stratagene) for further manipulation. Sequence analysis of DNA was done by Cetus. A genomic library of *Gossypium barbadense* L. cv. Sea Island was prepared by Clontech in phage vector EMBL 3. Propagation, screening, and isolation of phages were as described by Ausubel *et al.* (10). RNA was size-fractionated by electrophoresis on denaturing formaldehyde/agarose gels (11). Nick translation of DNA fragments was carried out using a kit from BRL and [^{32}P]dCTP (800 Ci/mmol; 1 Ci = 37 GBq; DuPont). Hybrid-selected RNA was translated in rabbit reticulocyte lysate (Promega) in the presence of [^{35}S]methionine (600 Ci/mmol, DuPont), and translation products were displayed on NaDodSO$_4$/PAGE gels (11).

Immunodetection of E6 Protein. Antibody was raised in rabbits against a 21-mer synthetic peptide (TTQK-PEEQEPRFIPETQNGYG) deduced from the nucleotide-derived amino acid sequence (Immuno Dynamics, La Jolla, CA). Immunoprecipitation and Western blot analysis of fiber proteins were carried out as described (11, 12).

Analysis of E6 Promoter in Transgenic Cotton. The coding region of the E6 gene in pSKSIE6-3B was replaced by a carrot extensin gene (13) (a gift from J. E. Varner, Washington University, St. Louis) as an *Nco* I–*Bst*XI fragment. The *Nco* I site in carrot extensin gene was introduced by a PCR using an *Nco* I site-containing primer (10). A marker gene, β-glucuronidase (*Gus*) with a cauliflower mosaic virus 35S promoter (14), was introduced into the chimeric E6–extensin construct at the *Not* I site. Stable transgenic Delta Pine 50 cotton plants were produced by introducing E6–extensin construct into cotton hypocotyl through particle bombardment and regenerating the plants (D. E. McCabe, personal communication).

RESULTS

The differential screening of the fiber cDNA library was designed to identify mRNAs that are preferentially expressed in fiber. Accordingly, the library was probed with cDNAs of 15- and 24-DPA fibers and leaf and 0-DPA ovule, flower, and root poly(A)$^+$ RNAs, and five putative fiber-specific cDNA clones were identified. All five clones showed preferential hybridization to fiber mRNAs in Northern blot analysis. To investigate whether any of the clones were expressed only at a given developmental stage, RNA blot analysis was carried out using mRNA from fibers of various developmental stages. RNA corresponding to each clone was detected throughout the development of the fiber (0- to 28-DPA fibers tested), although the profile of RNA concentration relative to the age of the fiber varied from clone to clone (data not shown). Detailed characterization of one of the clones, pCKE6, is described below.

Tissue and Developmental Regulation of E6 mRNA Expression. An RNA blot containing poly(A)$^+$ RNA (0.5 μg) from various cotton tissues was hybridized to the insert of pCKE6 (Fig. 1*A*). A broad hybridization signal was seen for 15- and 24-DPA fiber mRNAs, indicating multiple transcripts [1 and 1.1 kilobases (kb)]. On longer exposure, ovule, leaf, and flower tissue mRNA lanes also showed weak hybridization signals (Fig. 1*B*). RNA from 28-DPA fibers also contained E6 transcripts, whereas pollen and hypocotyl tissue RNAs showed no hybridization (data not shown). The size heterogeneity of RNAs may be due to differences in poly(A) length, transcription from different genes, or differential splicing

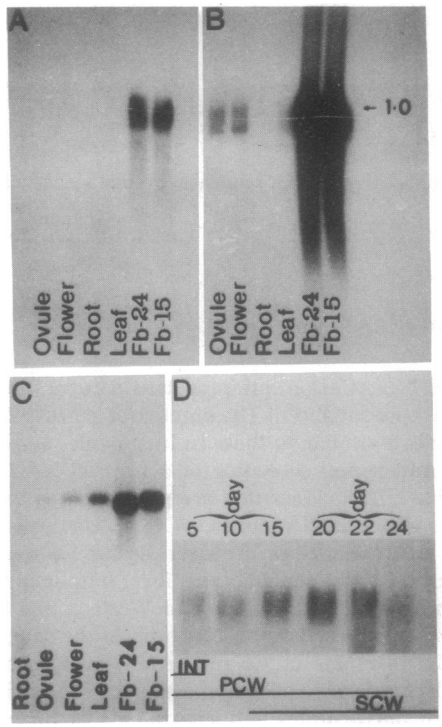

FIG. 1. Tissue and temporal expression of E6. (*A*) Poly(A)$^+$ RNA (0.5 μg) from each tissue was electrophoresed on formaldehyde/agarose gels, blotted to nitrocellulose, and hybridized to a ^{32}P-labeled E6 cDNA insert (1×10^8 cpm/μg:5.0×10^5 cpm/ml) for 12 hr. Filters were washed in buffers with various concentrations of SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0) that contained 1 mM EDTA and 0.2% NaDodSO$_4$ at 53°C for 20-min periods as follows: two washes in 2× SSC, two washes in 1× SSC, two washes in 0.5× SSC, and a final wash in 0.1× SSC. A 12-hr exposure of the autoradiogram is shown. Fb-15 and Fb-24 refer to RNA from 15- and 24-DPA fibers, respectively. (*B*) A 48-hr exposure of the autoradiogram in *A*. Multiple transcripts are discernable in ovule and flower RNA lanes. Molecular mass estimated from a *Hin*dIII digest of λ DNA is shown (1 kb). (*C*) To ascertain that the absence of hybridization in some lanes and the diffused nature of hybridization in other lanes in *A* were not due to RNA degradation, 0.5 μg of each mRNA was hybridized to clone pCKB11. pCKB11 is a cDNA clone that hybridized to various tissue RNAs. Hybridization to all tissue RNAs except root is observed. (*D*) Northern blot containing total RNA (20 μg) from 5- through 24-DPA fibers was hybridized to the insert of pCKE6. The 5-day RNA sample is from combined ovule and fibers, and all others are fiber RNAs only. Approximate time periods of fiber developmental stages are shown. INT, initiation; PCW, primary cell wall synthesis; SCW, secondary cell wall synthesis.

events but not from any degradation of RNAs as seen from the hybridization of a different clone, pCKB11 (Fig. 1*C*). On the day of anthesis, E6 transcript concentration was low in the ovules (Fig. 1 *A* and *B*). The RNA concentration increased as fiber development proceeded through initiation and primary cell wall synthesis, and decreased during the later stages of secondary cell wall deposition (Fig. 1*D*). The maximum concentration of E6 mRNA was observed in fibers of 15–22 DPA. The approximate concentration of E6 transcripts in fiber RNA was estimated by comparison of hybridization to *in vitro*-synthesized E6 transcripts of known specific activity. These studies showed that 20 μg of 10-, 15-, and 20-DPA fiber RNAs contained 1.7, 2.5, and 3.5 ng of E6 RNA, respectively (data not shown). Based on our observation that 20-DPA fiber RNA contained 1% poly(A)$^+$ RNA, ≈1.7% of poly(A)$^+$ RNA in a 20-DPA fiber is E6 transcript.

Structural Characterization. The complete nucleotide sequences of the pCKE6 insert (983 base pairs) and a cross-hybridizing clone pCKB3 were obtained. The sequence of

pCKB3 was identical to that of pCKE6 except for the 3′ end where 84 additional residues and a stretch of poly(A) were found. Fig. 2 shows the combined sequence of pCKE6 and pCKB3 (1067 base pairs). The cDNA insert contains an open reading frame 747 bases long with AUG codons at nucleotides 34, 61, and 94. If translation initiation occurs at each of these sites, it will result in peptides of 238 amino acids (29 kDa), 229 amino acids (27.5 kDa), and 218 amino acids (26 kDa), respectively. To confirm that E6 transcripts could be translated into discrete peptides and to assess their sizes, hybrid-selection RNA translation was carried out. Two peptides in the molecular mass range of 30–32 kDa were observed (Fig. 3A). The 238-amino acid peptide has a predicted isoelectric point of 4.99 and shows no N-linked glycosylation sites. Evaluation of hydrophilicity and hydrophobicity of E6 protein to predict membrane-buried and exposed regions indicated a highly hydrophobic N-terminal region (23 amino acids). The remainder of the peptide is hydrophilic in nature. Nucleotide sequence comparison of E6 with other gene sequences from GenBank (December 1991) revealed no significant homology. The nucleotide-derived amino acid sequence of E6 was also analyzed for various sequence patterns (motifs) that might suggest function of the E6 protein in the fiber. Motifs for three types of posttranslational phosphorylation [namely, casein kinase II (six sites), kinase C (eight sites), and tyrosine kinase (two sites)] were found (Fig. 2).

```
                                                    ↓
ACACACACAAGTAAA GCATTAGCAACCATA GCCATGGCTTCCTCA 45
                                      MetAlaSerSer
                                                  △
                ↓
CCAAAACTCTTCTCT ATGTCTATCCTCTTC CTTTTTGCCCTCTTC 90
ProLysLeuPheSer MetSerIleLeuPhe LeuPheAlaLeuPhe
↓
TCCATGCAAATCCAT GCTAGAGAGTACTTC AGCAAATTCCCAAGA 135
SerMetGlnIleHis AlaArgGluTyrPhe SerLysPheProArg

GTTAACATCAATGAG AAAGAGACAACAACC AGAGAGCAAAAGCAC 180
ValAsnIleAsnGlu LysGluThrThrThr ArgGluGlnLysHis
                  ▲△
GAGACCTTCGTTCCC CAGACCACCCAAAAG CCAGAAGAACAAGAG 225
GluThrPheValPro GlnThrThrGlnLys ProGluGluGlnGlu
                  △
CCAAGGTTCATTCCT GAAACCCAAAATGGT TATGGCCTTTACGGC 270
ProArgPheIlePro GluThrGlnAsnGly TyrGlyLeuTyrGly

CACGAGTCAGGCTCA AGCCGGCCCAGTTTC ACCACCAAAGAAACC 315
HisGluSerGlySer SerArgProSerPhe ThrThrLysGluThr
          △                     ▲△
TATGAACCCTATGTC ACCCCTGTTAGATTC CACCCTGATGAGCCC 360
TyrGluProTyrVal ThrProValArgPhe HisProAspGluPro

TATAACAGCATCCCC GAATCCTCCAACAAT AAAGACACTTACTAC 405
TyrAsnSerIlePro GluSerSerAsnAsn LysAspThrTyrTyr
●  ▲
TACAACAAGAATGCC TACGAGTCCACTAAG CAGCAAAACTTGGGC 450
TyrAsnLysAsnAla TyrGluSerThrLys GlnGlnAsnLeuGly
                  △
GAGGCCATTTTCACC GAGAAAGGATGGAGC ACCAAGGAAAACCAG 495
GluAlaIlePheThr GluLysGlyTrpSer ThrLysGluAsnGln
          △                     △
AACAACAACTACTAC AACGGCAACAATGGT TACAACAATGGCGAG 540
AsnAsnAsnTyrTyr AsnGlyAsnAsnGly TyrAsnAsnGlyGlu

AAGCAAGGCATGAGC GATACTAGGTACTTG GAGAATGGAAAGTAC 585
LysGlnGlyMetSer AspThrArgTyrLeu GluAsnGlyLysTyr
                                                ●
TACTATGACGTCAAG AGTGAGAACAACTAT TATCCAAACCGGTTC 630
TyrTyrAspValLys SerGluAsnAsnTyr TyrProAsnArgPhe

GACAACTCAAGAGGA GTTGCTTCGAGGAAC GAGTTCAATGAGAAT 675
AspAsnSerArgGly ValAlaSerArgAsn GluPheAsnGluAsn
                  ▲
CGTTACAACAACATG GGAAGGTACCACCAG AACCAAGAGGAGTTC 720
ArgTyrAsnAsnMet GlyArgTyrHisGln AsnGlnGluGluPhe

GAGGAAAGCGAGGAA GAGTTCGAACCCTGA TCACCTGTCGTACAG 765
GluGluSerGluGlu GluPheGluProEnd
    ▲
TATTTCTACATTTGA TGTGTGATTTGTGAA GAACATCAAACAAAA 810
CAAGCACTGGCTTTA ATATGATGATAAGTA TTATGGTAATTAATT 855
AATTGGCAAAAACAA CAATGAAGCTAAAAT TTTATTTATTGAGCC 900
TTGCGGTTAATTTCT TGTGATGATCTTTTT TTTTATTTTCTAATT 945
ATATATAGTTTCCTT TGCTTTGAAATGCTA AAGGTTTGAGAGAGT 990
TATGTTCTTTTTCTC TTCCTCTTTCTTTTT TAACTTTATCAAACA 1035
ATTTTTGAATAAAAA TGTGAGTATATTGTA AC 1067
```

FIG. 2. Combined nucleotide and deduced amino acid sequences of pCKE6 and pCKB3. The nucleotide-derived amino acid sequence is especially rich in glutamic acid (13 mol %) and asparagine (12.2 mol %). The putative poly(A) addition signal is underlined. The AUG codons near the 5′ end are marked by downward arrows (↓). Potential casein kinase II phosphorylation sites are identified by solid triangles, and amino acid residues phosphorylated by protein kinase C are marked by open triangles. Tyrosine kinase phosphorylation sites are indicated by solid circles. The sequence that was used for synthesis of synthetic peptide is underlined with a dashed line.

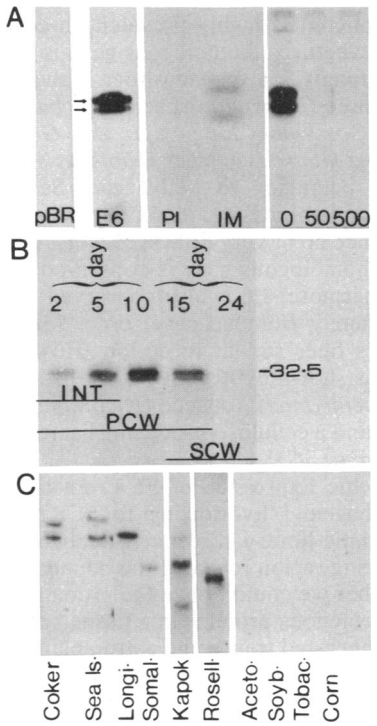

FIG. 3. Immunodetection of E6 protein. (A) pCKE6 was used to hybrid-select fiber poly(A)$^+$ RNA, which was then translated, and the radiolabeled proteins were electrophoresed on a NaDodSO$_4$/15% polyacrylamide gel (lane E6). pBR322 DNA was used as a negative control for the hybrid selection (lane pBR). Antibodies generated by immunizing rabbits with a synthetic E6 peptide were used to immunoprecipitate E6 proteins. Preimmune (5 μl, lane PI) and immune (2.5 μl, lane IM) sera were incubated with [^{35}S]methionine-labeled *in vitro* translation products (1 × 10^6 cpm) from fiber RNAs. Similarly, immunoprecipitations were repeated after addition of 0, 50, and 500 ng of synthetic E6 peptide (lanes 0, 50, and 500). (B) Proteins (75 μg) from 2- through 24-DPA fibers were subjected to NaDodSO$_4$/PAGE in 12% gels, transferred to nitrocellulose, and localized by immunostaining. The approximate times of initiation (INT), primary cell wall synthesis (PCW), and secondary cell wall synthesis (SCW) are shown. (C) Detection of homologous E6 genes in plants. Genomic DNAs (20 μg) from various plants (Sea Is, Sea Island cotton; Longi, *Gossypium longicalyx*; Somal, *Gossypium somalense*; Rosell, *Hibiscus sabdariffia* L. cv. Roselle; Soyb, soybean; and Tobac, tobacco) and cellulose-producing bacterium (Aceto, *A. xylinum*) were digested with *Eco*RI, blotted to nitrocellulose, and hybridized to the insert of pCKE6. The blot containing *A. xylinum*, soybean, tobacco, and corn was washed at low stringency (1× SSC, 53°C); other blots were washed at higher stringency (0.5× SSC, 53°C).

Immunodetection of E6 Protein. We tested whether antibodies against a synthetic 21-mer peptide of E6 can immunoprecipitate any proteins from *in vitro*-translated fiber RNAs. The 21-mer peptide was synthesized based on the nucleotide-derived amino acid sequence. Immunoprecipitation using E6 antiserum of translation products resulted in two proteins of ≈30 and 32 kDa (Fig. 3*A*). The observation that addition of synthetic peptide (50 ng–5 μg) to the translation mixture before immunoprecipitation resulted in the disappearance of these bands suggests that there are similar antigenic sites in the peptide and the immunoprecipitated proteins of the fiber (Fig. 3*A*). Furthermore, to confirm that the E6 antibody can cross-react with *in vivo* fiber protein, Western blot analysis was carried out with proteins isolated from fibers. Two proteins (30–32 kDa) were visualized (Fig. 3*B*). Thus these results confirm that the primary structure of E6 protein in fiber is likely to be identical to the deduced sequence shown in Fig. 2. E6 protein accumulation during fiber development was examined by Western blot analysis and found to be maximal between 5 and 15 DPA (Fig. 3*B*).

Since E6 mRNA is highly prevalent in cotton fibers, we investigated whether a homologous gene is present in other members of family Malvaceae (order Malvales). We tested DNA from three *Gossypium* species (*G. barbadense* L. cv. Sea Island, *Gossypium longicalyx*, and *Gossypium somalense*) and one *Hibiscus* species (*Hibiscus sabdariffa* L. cv. Roselle), all belonging to Malvaceae. Sea Island cotton produces fiber, whereas *G. longicalyx*, *G. somalense*, and Roselle produce no or very little seed hair. As seen from Fig. 3*C*, E6 or a homologous gene is conserved in all the above plants. Furthermore, E6 is also conserved in kapok (*Ceiba pentandra*, family Bombacaceae, order Malvales), a plant that produces fiber similar to cotton. However, E6 is not ubiquitous, as shown by negative hybridization results with soybean (*Glycine max*), tobacco (*Nicotiana tobacum*), corn (*Zea mays*), and a cellulose-producing bacterium, *Acetobacter xylinum* (Fig. 3*C*).

Tissue-Specific Expression of E6 Promoter in Transgenic Plants. A phagemid hybridizing to pCKE6 was isolated through genomic library screening. Sequence analysis of a 2-kb hybridizing region revealed the boundaries of E6 gene. To test whether we could utilize the promoter of E6 gene to express heterologous proteins in a tissue-specific manner in cotton, we generated transgenic cotton plants containing E6 promoter fused with a carrot extensin gene (13). Homologous sequences to carrot extensin were not detected in cotton tissues, including fiber (Fig. 4*A*). A 2.5-kb DNA fragment containing upstream sequences from the first initiation codon of E6 gene (pSKSIE6-3B) was fused to the coding region of the carrot extensin gene. This chimeric gene, when expressed, should result in a 1.5-kb transcript. Transgenic cotton plants were generated by particle bombardment and selection accomplished by *Gus* assays (D. E. McCabe, personal communication). Examination of fiber RNAs from five transformants showed expression of the correct size of carrot extensin RNA in fibers of three plants (Fig. 4*C*). Clearly the extensin transcript levels differed in each of the transformants. This is likely to be a result of the position at which the chimeric gene is integrated into the cotton genome. Southern blot analysis of four transgenic plants showed that all contained an integrated extensin gene (Fig. 4*F*). Various tissues from transgenic plant 6804 were examined for the presence of extensin RNA by Northern blot analysis. As Fig. 4*D* shows, only the fiber contained extensin RNA. When the same Northern blot was rehybridized to a control clone, pCKG8, hybridization was observed in leaf, ovule, fiber, and petals to prove RNA integrity in these samples (Fig. 4*E*). Examination of RNA from 5-, 10-, 15-, 20-, and 24-DPA fibers showed maximal extensin RNA concentration in 15- through 20-DPA fibers (data not shown). This pattern is similar to that of E6 expression (Fig. 1*D*).

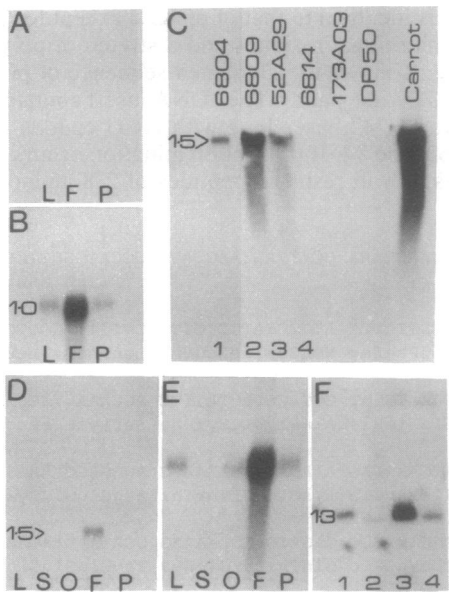

FIG. 4. E6 promoter-directed expression of extensin in transgenic cotton. (*A*) Northern blot containing total RNA (20 μg) from leaf (L), 15-DPA fibers (F), and petals (P) was hybridized to carrot extensin gene and washed in 1× SSC at 53°C. Absence of hybridization indicates that homologous sequences to carrot extensin are not present in these RNA samples. (*B*) The same blot shown in *A* was then hybridized to a control clone, pCKG8, to confirm presence of RNA in lanes L, F, and P. (*C*) Northern blots containing 20 μg of RNA from 15-DPA fibers of five transgenic plants (6804, 6809, 52A29, 6814, and 173A03), control plant, Delta Pine 50 (DP50), and carrot root were hybridized to extensin gene. Blots were washed in 0.5× SSC at 53°C. The carrot RNA was isolated from wounded carrot root as described (13). (*D*) Analysis of tissue-specific expression of carrot extensin in plant 6804. Northern blot containing poly(A) RNA (1 μg) from leaf (L), stem (S), ovule (O), fiber (F), and petal (P) was hybridized to the extensin gene and washed in 0.5× SSC at 53°C. Transcripts are present only in fiber. (*E*) Samples of the same RNAs shown in *D* were then hybridized to control clone pCKG8 to confirm hybridization to other RNAs. (*F*) Southern blot hybridization of DNA from four of the transgenic plants (*C* and *F*, lanes 1–4) showed an expected 1.3-kb band. The DNAs were digested with *Nco* I/*Xba* I. DNA from control plant Delta Pine 50 showed no hybridization to extensin gene (data not shown).

DISCUSSION

Our analysis of *in vitro*-translated proteins from various tissues of cotton fiber, ovule, and leaf by two-dimensional gel electrophoresis indicated that the number of genes that are active in fiber cells is very similar to those in tissues of multiple cell types (7). This result was also borne out by differential screening of the fiber cDNA library, where a majority of the fiber cDNA clones hybridized to different tissue mRNAs. Earlier studies in other plants have also shown examples where complexity of gene expression was similar in specialized and complex tissue types (15). Thus fiber cells, though devoid of any critical functions in the cotton plant, contain large numbers of active genes common to other cell types along with fewer active genes unique to fiber. A second point that became clear during the cDNA library screening was that no major changes occur in the mRNA population during primary and secondary cell wall synthesis stages. We did not detect any subset of genes that are exclusively expressed during a given developmental stage. Thus, it seems likely that most of the genes in fiber are active throughout the development of the fiber. Alterna-

tively, gene transcription ceases early in the fiber development, but differential mRNA utilization occurs during growth. Although it is clear that no gross changes occur in the RNA population, the protein content may change during fiber development. Earlier measurement of protein content in the cotton cell wall (weight percent) has shown it to decrease from a high of 40% in 5-DPA fibers to <2% in 18-DPA fibers (16). Measured as weight per unit length, the fiber wall protein content peaks in ≈16-DPA fibers before rapidly declining (16). The level of steady-state E6 RNA was high in 20-DPA fibers and persisted beyond 24-DPA fibers. However, there was a sharp decline of E6 proteins in 15-DPA and older fibers. Thus, the decline of E6 protein in 15-DPA fibers and the near absence of E6 protein in 24-DPA fibers (shown in this study) can be interpreted to be due to a protein degradation occurring in fiber cells during this time period. This degradation must be selective since a major biosynthetic activity, cellulose synthesis/deposition, occurs in fibers 24 DPA and older, and protein components necessary for this event would, therefore, be expected to be preserved.

Though the principal component of cotton fiber is cellulose, it is a composite of many chemical constituents whose interactions and arrangement in the cell form the basis of fiber properties. Genes that are preferentially expressed in fiber may function in the synthesis/deposition of these components. Thus, the five cDNA clones identified during the screening are likely to be important. To our knowledge, cotton E6 is one of the first fiber protein genes identified and characterized. It contains several sequence motifs that indicate potential posttranslational modifications. Phosphorylation sites for three kinases have been detected. The possibility that E6 protein may be a substrate for casein kinase II is intriguing. The known physiological substrates of casein kinase II include proteins involved in transcription and translation, regulatory proteins, enzymes of metabolic pathways, and structural proteins (for a review, see ref. 17). Based on the abundance of mRNA and protein and the predominantly tissue-specific and developmentally regulated expression of E6, a general transcription/translation or regulatory function is unlikely. A direct role in the cellulose synthesis can also be ruled out since the maximum rate of cellulose deposition in fibers occurs between 26 and 28 DPA (16), when E6 protein concentration is very low. However, a number of other carbohydrate components such as neutral sugars, uronic acid, and its polymer, pectins are synthesized early in the fiber development. Thus, E6 could be a candidate for the biosynthesis or degradation of any of these polysaccharides. Alternatively, E6 may have a structural role in fiber. A number of models based on chemical and electron microscopical studies of algal cell walls, which are considered to be models for cotton fiber, predict structural roles for proteins in the architecture of cell walls where cellulose microfibrils are embedded in a matrix consisting of hemicellulose, pectins, proteins, and other unidentified components that interact with each other‡ (refs. 18 and 19 and references therein). If E6 protein fulfills such a structural role in fiber, the proline (after hydroxylation) and serine residues would be candidates for interaction with carbohydrate moieties. Similarly, glutamyl and lysyl residues could form amide cross-links imparting stability. A unique structural role or enzymatic role in the primary cell wall deposition in fiber is in agreement with its predominant expression of E6 and its conservation in fiber-producing plants but not in all plants or in *A. xylinum*, where cellulose is not deposited in cell walls.

The regulatory elements of genes that are active in fiber are of special interest due to their potential in expressing heterologous genes in fiber. Such a capability is essential for the eventual modification of fiber to generate specialty fibers. Based on the RNA analysis, the E6 gene is most active in fiber, though low levels of transcripts are found in leaf, ovule, and flower. The hybridization to 0-DPA ovule and flower RNAs may be due to the presence of fiber initials destined to become fiber cells (20). It is not clear what cell types express E6 gene in leaf. E6 promoter is thus expected to direct expression of genes predominantly in fiber, a result confirmed by expression of carrot extensin gene in transgenic cotton. From these studies it can be concluded that regulatory elements responsible for tissue and developmental expression are located in the 2.5-kb promoter fragment we isolated. Therefore, this promoter will be a valuable tool in the study of gene expression in fiber.

We are indebted to Dr. Dennis McCabe for generating transgenic cotton plants. We thank Drs. William Swain and Winston Brill for useful discussions and critical comments and Greg Keller, Grant Johnson, and Michael Petersen for excellent technical assistance.

1. Basra, A. S. & Malik, C. P. (1984) *Int. Rev. Cytol.* **89**, 65–113.
2. Kosmidou-Dimitropoulou, K. (1986) in *Cotton Physiology*, eds. Mauney, J. R. & Stewart, J. M. (Cotton Found., Memphis, TN), Vol. 1, pp. 361–374.
3. Ramsey, J. C. & Berlin, J. D. (1976) *Bot. Gaz. (Chicago)* **137**, 11–19.
4. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G. A., Van Boom, J. H. & Benziman, M. (1987) *Nature (London)* **325**, 279–281.
5. Wong, H. C., Fear, A. L., Calhoon, R. D., Eichinger, G. H., Mayer, R., Amikam, D., Benziman, M., Gelfand, D. H., Meade, J. H., Emerick, A. W., Bruner, R., Ben-Bassat, A. & Tal, R. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8130–8134.
6. Seagull, R. W. (1986) *Can. J. Bot.* **64**, 1373–1381.
7. John, M. E. & Stewart, J. M. (1992) *Trends Biotechnol.* **10**, 165–170.
8. Dellaporta, S. L., Wood, J. & Hicks, J. B. (1984) *Molecular Biology of Plants: A Laboratory Course Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 36–37.
9. D'Alessio, J. M., Noon, M. C., Ley, H. L., III, & Gerard, G. F. (1987) *Focus* **9**, 1–4.
10. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) *Current Protocols in Molecular Biology* (Wiley, New York).
11. John, M. E., John, M. C., Ashley, P., MacDonald, R. J., Simpson, E. R. & Waterman, M. R. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5628–5632.
12. Blake, M. S., Johnston, K. H., Russell-Jones, G. J. & Gotschlich, E. C. (1984) *Anal. Biochem.* **136**, 175–179.
13. Chen, J. & Varner, J. E. (1985) *EMBO J.* **4**, 2145–2151.
14. McCabe, D. E., Swain, W. F., Martinell, B. J. & Christou, P. (1988) *Bio/Technology* **6**, 923–926.
15. Willing, R. P. & Mascarenhas, J. P. (1984) *Plant Physiol.* **75**, 865–868.
16. Meinert, M. C. & Delmer, D. P. (1977) *Plant Physiol.* **59**, 1088–1097.
17. Pinna, L. A. (1990) *Biochem. Biophys. Acta* **1054**, 267–284.
18. Preston, R. D. (1986) in *Cellulose, Structure, Modification, and Hydrolysis*, eds. Young, R. A. & Rowell, R. M. (Wiley, New York), pp. 3–27.
19. Bacic, A., Harris, P. J. & Stone, B. A. (1988) in *The Biochemistry of Plants*, eds. Stumpf, P. K. & Conn, E. E. (Academic, New York), Vol. 14, pp. 297–371.
20. Stewart, J. M. (1975) *Am. J. Bot.* **62**, 723–730.

‡Zahn, H., *Latest Findings on the Microstructure of Cotton*, Abstracts of the 19th International Cotton Test Conference, February 25–27, 1988, Fiber Institute, Bremen, F.R.G.