Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 5140–5144, August 1984
Cell Biology

A single rat fibronectin gene generates three different mRNAs by alternative splicing of a complex exon

(fibronectin/gene structure/splicing/RNA processing)

JOHN W. TAMKUN, JEAN E. SCHWARZBAUER, AND RICHARD O. HYNES

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Phillip A. Sharp, April 23, 1984

ABSTRACT **Three fibronectin mRNAs exist in rat liver, differing by the presence or absence of segments of 285 or 360 bases at a point within the coding region. We previously proposed that the three mRNAs are encoded by a single gene and arise via alternative splicing of a common transcript. In order to test this hypothesis, we have isolated clones spanning approximately half of the fibronectin gene from a Fisher rat genomic library; blot hybridization analyses reveal the presence of only one fibronectin gene in the haploid rat genome. We determined the sequence of a portion (1221 nucleotides) of this gene. This sequence shows clearly that the three fibronectin mRNAs encoded by this gene are generated by a pattern of alternative splicing in which one 5′ splice site can be paired with any one of three 3′ splice sites, one at the beginning of, and two within, a single complex exon.**

Fibronectins are high molecular weight glycoproteins that are involved in a wide range of biological processes including cellular adhesion, cytoskeletal organization, cell migration, hemostasis and thrombosis, and malignant transformation (1, 2). Fibronectins are composed of similar but nonidentical 210–250 kDa polypeptide subunits. Several distinct forms of fibronectin exist and have been studied in some detail. One form is present at considerable levels (300 μg/ml) in plasma and is synthesized and secreted by hepatocytes (3). A related fibronectin is found in fibrillar matrices at the surfaces of many cell types, such as fibroblasts and endothelial cells. Post-translational modifications alone probably cannot account for the differences between plasma and cellular fibronectins and among their respective subunits (1–5). This has led to speculation concerning the number of genes and mRNAs encoding fibronectins.

We recently reported the isolation and sequencing of several fibronectin cDNA clones from a rat liver library constructed in the expression vector λgt11 (6). S1 nuclease analysis demonstrated the existence of at least three different fibronectin mRNAs in rat liver. The mRNA sequences predicted from the cDNA sequence are identical except for a region 827 bases 5′ of the termination codon. One type of mRNA contains 285 nucleotides at this position, while a second contains this sequence and an additional 75 nucleotides 5′ to this segment, resulting in the introduction of a total of 360 nucleotides at this position. The third type of mRNA lacks the entire 360-base segment. The three types of mRNA thus encode fibronectin subunits varying by the presence or absence of stretches of 95 or 120 amino acids. These segments are located 276 residues from the carboxyl terminus of fibronectin, near the cell- and heparin-binding domains, and are likely to account for the difference between the two different subunits of plasma fibronectin. These segments also may be responsible for one or more properties of the different forms of fibronectin, possibly including self-association and interactions with other molecules or with cells (6). These possibilities are currently under investigation, but clearly a question of major interest concerns the origin of the different mRNAs.

The identity of nucleotide sequence on either side of the region of difference suggested that these mRNAs are encoded by a single gene (6). Previous Southern blotting analyses have failed to reveal the existence of more than one human or bovine fibronectin gene (7), but the limited nature of these analyses did not allow this possibility to be eliminated. If there were a single gene, the three mRNAs would have to arise by alternative splicing of a common transcript. Possibilities include separate exons encoding each of the segments and alternative splicing of a single exon. The presence of 3′ splice sites within the coding sequence led us to favor the second of these models (6), but we could not eliminate other possibilities. Obviously, these hypotheses can be tested only by examining the structure of the gene or genes encoding fibronectin.

In this paper, we report the isolation of overlapping fibronectin genomic clones from a rat library constructed in the vector EMBL 3B (8). Partial sequencing of these clones, as well as Southern blot analyses, show that a single fibronectin gene does indeed encode at least three different fibronectin mRNAs, and that two of these mRNAs arise through the utilization of 3′ splice sites contained within a single exon. This represents a pattern of alternative splicing not previously seen for other cellular genes.

MATERIALS AND METHODS

Materials. All enzymes were purchased from New England Biolabs, except where otherwise noted. Adult male Fisher rats were purchased from Charles River Laboratories. Bacteriophage λ EMBL 3B (8) DNA was generously provided by Norma Neff.

Preparation of Bacteriophage DNA. Phage (3.75×10^7) were adsorbed to 2×10^9 bacteria (*Escherichia coli* strain LE 392) in 10 mM $MgSO_4$ for 20 min at 37°C. NZCM medium (9) (50 ml) was added, and the culture was shaken vigorously at 37°C until lysis occurred (7 hr). Phage DNA was isolated as described for plate lysates (9). For large-scale (≤2 liters) preparation of phage DNA, the purification procedure of Yamamoto *et al.* (10) was used.

Preparation of Vector. EMBL 3B DNA was digested with a 5-fold excess of both *Bam*HI and *Eco*RI. After phenol extraction, the digested DNA was precipitated three times with 0.1 M NaCl and an equal volume of isopropanol to remove the short *Bam*HI/*Eco*RI-generated fragments from the *Bam*HI-generated arms and the *Eco*RI-generated internal fragment.

Preparation of Rat DNA. DNA was prepared from adult male Fisher rat livers as described by Blin and Stafford (11)

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.

Abbreviation: kb, kilobase.

except that the RNase digestion was omitted and equilibrium centrifugation in cesium chloride gradients was included as a final purification step (12). Rat DNA was partially digested with an amount of *Sau*3A I (0.17 units/μg) determined to yield the maximum proportion of 18- to 22-kilobase (kb) fragments (13). Digested DNA was sedimented through 10–40% linear sucrose gradients (9). Fractions containing fragments from 16 to 24 kb in size were pooled, dialyzed against 10 mM Tris chloride, pH 8.0/1 mM EDTA, and precipitated with 0.2 M NaCl and two volumes of ethanol.

Construction of Genomic Library. The 16- to 24-kb *Sau*3A I fragments of rat DNA were ligated to EMBL 3B arms for 16 hr at 15°C with T4 DNA ligase. Packaging extracts were prepared and used as described by Grosveld *et al.* (14). The highest cloning efficiencies (approximately 5×10^5 pfu/μg of rat DNA) were obtained by using a 3:1 molar ratio of arms to insert at an insert concentration of 42 μg/ml. Ligation and packaging of arms alone under identical conditions yielded low numbers of nonrecombinant phage (<3000 pfu/μg of ligated arms). Approximately 10^7 independent recombinant phage were produced and stored at −80°C in 7% dimethyl sulfoxide.

Isolation of Fibronectin Genomic Clones. The probes used for screening the genomic library were prepared from the rat liver fibronectin cDNA clone λrlf-2 (ref. 6; see Fig. 1). After digestion of λrlf-2 DNA with *Eco*RI, the two insert fragments were separated by agarose gel electrophoresis, eluted by the glass bead method (15), and labeled by nick-translation (16) to a specific activity of 10^8 cpm/μg. Independent recombinants (7.5×10^5) were screened by plaque hybridization (17) at a density of 150 plaques per cm^2. Positive signals were taken through an additional three rounds of screening at low-plaque density before plate stocks were made.

Southern Blot Analysis. Phage or genomic DNA was digested with restriction enzymes, fractionated by agarose gel electrophoresis, and transferred to Zetabind (AMF Cuno Division, Meriden, CT) essentially as described by Southern (18) with the following modification. After baking at 80–90°C for 1–3 hr in a vacuum oven, filters were washed for 1–2 hr at 65°C in 0.1× NaCl/Cit containing 0.1% NaDodSO$_4$ (1× NaCl/Cit = 0.15 M NaCl/0.015 M sodium citrate, pH 7). Filters were incubated for 3 hr at 65°C in 6× NaCl/Cit containing 0.5% NaDodSO$_4$, 5× Denhardt's solution (1× = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin) and 100 μg of *E. coli* DNA per ml prior to hybridization. Hybridization was allowed to proceed for 16 hr at 60°C in 6× NaCl/Cit containing 10 mM EDTA, 5× Denhardt's solution, 0.5% NaDodSO$_4$, 100 μg of *E. coli* DNA per ml, and 10% dextran sulfate with 10 ng of probe ($\approx 10^8$ cpm/μg) per filter. Filters were washed with several changes of either 2× NaCl/Cit containing 0.1% NaDodSO$_4$ at 50°C (low stringency) or 0.1× NaCl/Cit containing 0.1% NaDodSO$_4$ at 65°C (high stringency).

DNA Sequence Determination. End-labeled fragments were sequenced by the method of Maxam and Gilbert (19).

RESULTS

Isolation of Fibronectin Genomic Clones. Two different *Eco*RI fragments from the fibronectin cDNA clone λrlf-2 were used to screen the EMBL 3B rat genomic library (see Fig. 1). Probe A contains 583 bases of 3′ untranslated sequence and the last 310 bases of coding sequence. Probe B contains the preceding 1200 nucleotides of coding sequence, including the 360-nucleotide difference segment. Nine fibronectin genomic clones were isolated from approximately 750,000 independent recombinants screened. After partial restriction mapping and Southern blot analyses of all nine clones, two clones were selected for further study (Fig. 1).

Both phage contain 18-kb inserts and share 7 kb of overlapping sequence. Clone λrFN-1 hybridizes to both probes A and B, while clone λrFN-2 hybridizes to probe B alone. DNA sequencing revealed that the last *Eco*RI site in λrFN-1 encodes the *Eco*RI site in the untranslated region of the cDNA (data not shown). Since this *Eco*RI site is 2 kb from the 3′ end of the λrFN-1 insert, this clone probably includes the 3′ end of the gene. The chicken fibronectin gene has been reported to have an approximate length of 48 kb (20). Assuming that both rat clones are derived from the same gene, λrFN-1 and -2, which together span 29 kb of genomic DNA, represent over half of the rat fibronectin gene.

Southern Blot Analysis. To determine the number of fibronectin genes present in the rat genome, we wanted to use a probe that had a high probability of detecting multiple genes, should they exist. We were also particularly interested in the differences between mRNAs. Accordingly, we chose as a probe a *Bgl* II/*Mbo* II fragment from λrlf-2, which includes 100 nucleotides common to all three cDNAs and most of the region of difference (see Fig. 3). If the three different fibronectin mRNAs were not encoded by a single gene, one would expect this probe to yield multiple bands when used for Southern blots. Because multiple bands are often ob-

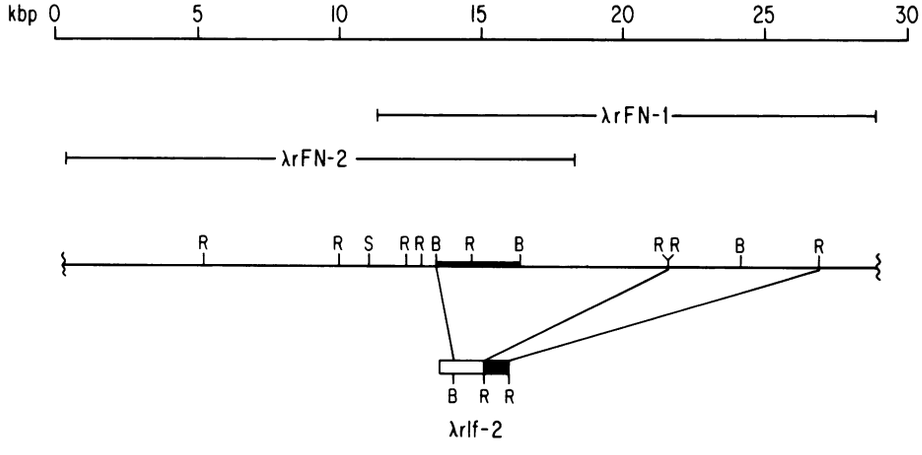

FIG. 1. Restriction enzyme map of the 3′ half of the rat fibronectin gene. The region was mapped by Southern blot analysis with probes derived from the fibronectin cDNA clones λrlf-1 and λrlf-2 (6) and from restriction mapping of the overlapping genomic clones λrFN-1 and λrFN-2. The λrlf-2-derived probes used for the screening of the genomic library are represented by shaded (probe A) and unshaded (probe B) boxes. The *Eco*RI site at the 5′ end of probe B does not exist in the cDNA but represents a synthetic *Eco*RI linker introduced in the construction of the cDNA library. The 3.2-kb *Bam*HI fragment (bold line) of λrFN-1 was subcloned in pBR322, mapped with restriction enzymes, and partially sequenced (see Fig. 3). R, *Eco*RI; S, *Sal* I; B, *Bam*HI. kbp, Kilobase pairs.

served on Southern blots due to interruptions of the genomic DNA by introns, we included digests of genomic clone DNA on our blots to aid in interpretation of the results.

A Southern blot of both Fisher rat genomic DNA and λrFN-1 DNA digested with a variety of enzymes is shown in Fig. 2. For all five enzymes used, a single band was found to hybridize with the *Bgl* II/*Mbo* II probe. The patterns observed with genomic and phage DNAs were identical. The intensity of the signals for the genomic DNA and the phage DNA, which was present at a level equivalent to one copy per haploid genome, were roughly equal. These results strongly suggest that the three mRNAs arise from a single gene, as predicted from the identity of nucleotide sequences surrounding the region of difference in the cDNAs. This experiment has been repeated with a number of probes covering the available 3 kb of fibronectin mRNA sequence (see ref. 6; data not shown). In all cases, the patterns observed with genomic DNA were consistent with those seen with λrFN-1 or -2 DNA. These Southern blotting results provide no evidence for the presence of more than one fibronectin gene in the rat genome and show that a single gene encodes the three different mRNAs. Therefore, these mRNAs must be the result of alternative splicing.

Sequence Determination Studies. What pattern of RNA splicing results in the formation of these three mRNAs from a common transcript? Southern blotting of restriction enzyme digests of λrFN-1 and λrFN-2 using cDNA probes covering the difference segment showed that this entire region was contained within a single 3.2-kb *Bam*HI fragment (see Fig. 1). A continuous sequence of 1221 nucleotides determined from this fragment (Fig. 3) contains sequence on both sides of and including the region of difference. The sequence

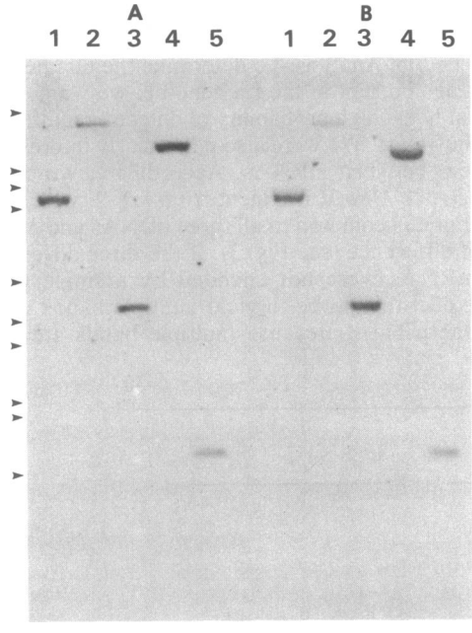

FIG. 2. Determination of fibronectin gene copy number by blot hybridization. λrFN-1 DNA (*A*) or Fisher rat genomic DNA (*B*) was digested with *Bgl* II (1), *Hin*dIII (2), *Eco*RI/*Bam*HI (3), *Eco*RI (4), or *Xba* I (5). Ten micrograms of each digested genomic DNA (0.14 ng of digested phage DNA) was fractionated on a 1.0% agarose gel. After transfer to Zetabind, filters were probed with the 401-bp *Bgl* II/*Mbo* II fragment of λrlf-2 (see Fig. 3). Low-stringency conditions were used for hybridization (6× NaCl/Cit at 60°C) and washing (2× NaCl/Cit at 50°C). Note the generation of single bands by each enzyme used and the identity of patterns observed for both genomic and phage DNAs, strongly suggesting the presence of a single fibronectin gene per haploid rat genome. The position of 21.8-, 5.24-, 4.21-, 3.41-, 1.98-, 1.57-, 1.32-, 0.93-, 0.84-, and 0.58-kb *Hin*dIII/*Eco*RI fragments of phage λ DNA are marked by arrowheads.

begins and ends in introns and contains two complete exons. The first exon contains 90 bases of coding sequence. The 3′ boundary of this exon coincides exactly with the last nucleotide of sequence common to all three fibronectin mRNAs. The intron that follows consists of 613 nucleotides of sequence, which contains multiple termination codons in all three reading frames. This intron has typical 5′ and 3′ splice sites (see D1 and A1 in Fig. 4). The organization of the following 467-base exon is quite interesting. The 5′ boundary of this exon is the first of the 360 nucleotides present in the largest of the mRNAs. The entire 360-base segment is then encoded without interruption by introns. This segment is not immediately followed by an intron but rather by 107 nucleotides of sequence common to all three forms of fibronectin mRNA. That is, coding sequences of the three different mRNAs, with and without inclusions of 285 or 360 bases, are all found within a single exon. As previously noted in the sequences of the cDNA clones, this segment of coding sequence contains two blocks of nucleotides that fit the consensus sequence for 3′ splice sites (see A2 and A3 in Fig. 4).

DISCUSSION

The results reported here clearly demonstrate that the three rat liver fibronectin mRNAs reported previously (6) are encoded by a single gene and must be formed by alternative processing of a common transcript. The sequence of the gene shows that a novel pattern of RNA splicing produces the different fibronectin mRNAs (Fig. 4). As predicted, one 5′ splice site (D1) can be paired with one of three different 3′ splice sites (A1–A3). To produce the mRNA containing the entire 360-base segment, the 3′ splice site (A1) at the end of the 613-nucleotide intron is chosen, resulting in the splicing out of this intron and the inclusion of the largest possible coding sequence. When either of the other two 3′ splice sites (A2 or A3) is used, the 5′ 75 or 360 nucleotides of the exon are deleted along with the intron, resulting in the generation of the two smaller mRNAs. Thus, two 3′ splice sites within the second exon can be used with the 5′ splice site at the end of the first exon to generate two additional types of fibronectin mRNA.

Could other 3′ splice sites be utilized by the 5′ splice site of the 613-base intron in order to create yet other fibronectin mRNAs? The presence of a termination codon (TAA) immediately preceding the 3′ splice site of this intron precludes the possibility that the exon could extend further in the 5′ direction and encode insertions of coding sequence longer than 360 bases. Furthermore, while other potential 3′ splice sites exist in the 467-nucleotide exon, use of any of these sites together with D1 would result in the introduction of a termination codon into the reading frame. Finally, we have examined the sequence of the 613-base intron for potential extra exons flanked by acceptor and donor splice sites. There is such a short block of sequence, starting at either position 265 or position 289 and ending at position 322 (see Fig. 4), which could encode 12 or 20 amino acids. Combination of the 5′ splice site at the end of this sequence with the 3′ splice sites (A1–A3) of the 467-base exon would give rise to termination codons. However, this 5′ splice site could be paired with a potential 3′ splice site at position 951, which would lead into the same reading frame as does the combination of D1 with A1–A3. We previously have noted the possible existence of other mRNAs; S1 nuclease analysis gave a minor band that could indicate the presence of an mRNA arising from a splice at position 951 (see ref. 6). Therefore, while we only have direct evidence for generation of three mRNAs by alternative splicing of this complex region of the rat fibronectin gene, we cannot rule out the possibility of a fourth, although this would require yet another pattern of alternative splicing.

```
                                        32                                  62
GTTTCCCATCA GGT CTG GAG CCA GGA ACC GAG TAC ACC ATC TAT GTC ATC GCA CTG AAG AAC
            GLY LEU GLU PRO GLY THR GLU TYR THR ILE TYR VAL ILE ALA LEU LYS ASN

                                        92                                 121
AAT CAG AAG AGT GAG CCC CTG ATT GGG AGG AAA AAG ACA G G TAA AGA CTC CTA ACC GAC
ASN GLN LYS SER GLU PRO LEU ILE GLY ARG LYS LYS THR

                                       151                                 181
CGT TGG TTT GCA GGG TGA AGT TCT CTA CGT TTT GAC ACA TCG CTC CTG GAT TTT TTT TTT

                                       211                                 241
TCT AGC AGA ATG GCT TAT ATA TTG TAC TAT TAG GCA CAA ATC ACC TAA CAT GCT CTA AAA

                                       271                                 301
ATG TCT TAT TTA TTT TTC TAC AGC CTT AAG CTT CTT TTA CCT TGC AGA AAT TCA AAT TAG

                                       331                                 361
ATT GTT TGA ATT GGG GTT GAG GTG GGG GGG GGC TTT TTA ATG GTG AGA GTA AAA CCT CTC

                                       391                                 421
CGA TTT GTG TTG ATT TGT TTA TTA TTT GAT TTC TGA ACT TCA GAT ATT TCA GTT AGT TGC

                                       451                                 481
AAA GGC TTA ACT GGA GCA GAG ATG GCT GTC TGT GGG GCT CAG CGG TTC AAA GCT TTG TCT

                                       511                                 541
GTA TGA GAA TTC AAA CTA ACA TTT TTT CCC CCA TCT AAA CTA TAA AGA AAT CTG TGA AAA

                                       571                                 601
GAA GAA AAA ATC CAG ATT TCT TTC TAG AGT GTT CTG TGA TCT GAG AAC CGC TGG CTG TCT

                                       631                                 661
GGG AGC ATT GAT CGG GTT GAT ATA CAG TTG TTG CCT GGA TTG ATG GTG ATT GGT TGC TTC

                                       691                                 720
TTC TTT GAG CTT AAC GCG CTT TGC TTT TTT GGC TCT AAC CTC TCT CTT GGC TAG AT GAG
                                                                      ASP GLU

                                       750                                 780
CTT CCC CAA CTG GTT ACC CTT CCA CAC CCC AAT CTT CAT GGA CCA GAG ATC TTG GAT GTT
LEU PRO GLN LEU VAL THR LEU PRO HIS PRO ASN LEU HIS GLY PRO GLU ILE LEU ASP VAL

                                       810                                 840
CCC TCC ACA GTT CAA AAG ACC CCC TTC GTC ACC AAC CCT GGG TAT GAC ACC GAA AAT GGT
PRO SER THR VAL GLN LYS THR PRO PHE VAL THR ASN PRO GLY TYR ASP THR GLU ASN GLY

                                       870                                 900
ATT CAG CTT CCT GGC ACA TCC CAC CAA CAA CCC AGT GTT GGG CAA CAA ATG ATC TTT GAG
ILE GLN LEU PRO GLY THR SER HIS GLN GLN PRO SER VAL GLY GLN GLN MET ILE PHE GLU

                                       930                                 960
GAA CAT GGC TTT AGG CGA ACC ACG CCA CCC ACT GCG GCC ACC CCC GTC AGG CTT AGG CCA
GLU HIS GLY PHE ARG ARG THR THR PRO PRO THR ALA ALA THR PRO VAL ARG LEU ARG PRO

                                       990                                1020
AGA CCA TAC CTG CCG AAT GTA GAT GAG GAG GTC CAA ATC GGT CAT GTT CCC AGG GGA GAC
ARG PRO TYR LEU PRO ASN VAL ASP GLU GLU VAL GLN ILE GLY HIS VAL PRO ARG GLY ASP

                                      1050                                1080
GTA GAC TAC CAC CTC TAT CCT CAT GTT CCG GGG CTC AAT CCA AAT GCC TCT ACA GGA CAA
VAL ASP TYR HIS LEU TYR PRO HIS VAL PRO GLY LEU ASN PRO ASN ALA SER THR GLY GLN

                                      1110                                1140
GAA GCT CTC TCT CAG ACA ACC ATC TCT TGG ACG CCA TTC CAG GAG AGT TCT GAG TAC ATC
GLU ALA LEU SER GLN THR THR ILE SER TRP THR PRO PHE GLN GLU SER SER GLU TYR ILE

                                      1170                                1200
ATT TCA TGC CAA CCT GTT GGC ACT GAC GAA GAG CCC TTA CAG GTA TAT ATT ACA GCC CTT
ILE SER CYS GLN PRO VAL GLY THR ASP GLU GLU PRO LEU GLN

ATT ATG AGT GTC TCT CTA GAT
```

FIG. 3. Nucleotide sequence of the portion of the rat fibronectin gene encoding the region of difference between the fibronectin mRNAs. The 1221 nucleotides of continuous sequence were obtained from the 3.2-kb *Bam*HI fragment of λrFN-1 (see Fig. 1). The amino acid sequence encoded by the two exons present in this sequence are shown. Exon/intron boundaries are marked by arrowheads. The positions of the *Bgl* II (A-G-A-T-C-T) and *Mbo* II (G-A-A-G-A) sites used for generating the probe used for the Southern blot (see Fig. 2) are underlined.

RNA transcripts of several cellular genes are known to be differentially processed to produce mRNAs encoding variant polypeptides. Alternative splicing can result in the deletion of entire exons, as in the case of *Drosophila* myosin (23) and murine αA-crystallin (24). The use of alternative polyadenylation sites permits the deletion of exons at the 3′ end of RNA transcripts, allowing single genes to encode both the secreted and membrane-associated forms of immunoglobulins (25, 26). The failure to splice out an intron creates mRNAs encoding rat fibrinogen chains (γA and γB) with different COOH-termini (27). The use of alternative promoters in the yeast invertase (28), murine α-amylase (29) and myosin light chain (30) genes provides yet another mechanism for the production of multiple mRNAs from a single gene. However, we know of no other cellular gene in which a single 5′ splice site can combine with several 3′ splice sites as in the case of fibronectin. The closest parallels occur in animal viruses such as adenovirus, polyoma, and simian virus 40 (31).

Although the data presented show clearly that the three fibronectin mRNAs arise from a single gene and, further-

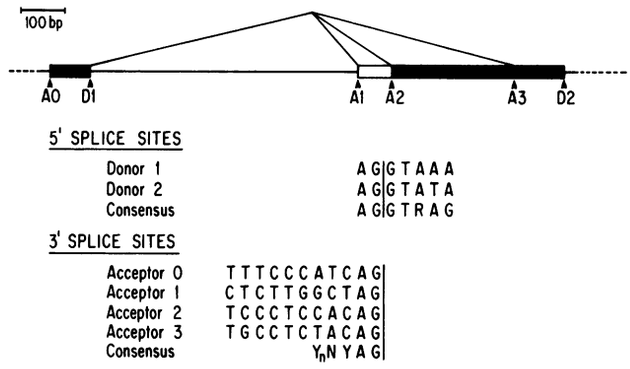

FIG. 4. Alternative splicing pattern giving rise to the three fibronectin mRNAs found in rat liver. Introns are represented by thin lines. Regions encoding sequence common to all three forms of fibronectin mRNA are shown as black boxes. The 285- and 75-nucleotide-included segments are shown as crosshatched and open boxes, respectively. Splice donor (D) and acceptor (A) sites are marked by arrowheads. Below this, the sequences of these sites are shown and compared with consensus sequences for 5′ and 3′ splice sites (32, 33) in which R = purine, Y = pyrimidine, and N = any base. bp, Base pairs.

more, that there are no other closely related genes, it is important to consider whether or not the data rule out the possibility of other fibronectin genes. We have detected none by Southern blotting analyses covering about 30% of the coding sequence (Fig. 2; unpublished data) using conditions of moderately low stringency which would have detected genes containing significantly diverged homologous sequences. However, it is known that the sequence of fibronectin contains homologous internal repeats of three different types (32) that, while clearly homologous at the protein level, differ by more than 50% at the level of nucleotide sequence (6, 33). The conditions used for the Southern blotting experiments do not reveal crossreactions between homologous repeats (type III) that are present in the region analyzed. Therefore, we would not have detected other genes that are as closely related as are the endoduplications within one fibronectin gene. If the stringency is lowered further, nonspecific reactions begin to produce unacceptable background signals. Therefore, we can conclude that, within the range of homology detectable by Southern blotting analyses, there is only a single fibronectin gene; but we cannot rule out the presence of genes encoding related proteins. Nevertheless, it is clear that the known fibronectin mRNA sequences are accounted for by a single gene.

Kornblihtt *et al.* (33) have reported the existence in a cultured cell line of human fibronectin mRNAs differing by the presence or absence of 270 nucleotides of sequence encoding an additional 90-amino-acid type III homology repeat. This sequence is located 813 nucleotides 5′ of the region of difference found in the rat fibronectin mRNAs. Perhaps significantly, the difference found between the human mRNAs occurs at the end of a type III homology repeat (33), as do the differences present in the rat mRNAs (6). However, the inserted segment in the rat case is not a type III repeat. Variation among human mRNAs at the position of difference found in the rat mRNAs was not investigated for the human cDNA clones, and we have not found any evidence for the existence of the segment encoding the additional type III repeat in any of the rat liver cDNA clones analyzed. However, if both alternatives can occur in both species, the potential exists for the production from one gene of six fibronectin mRNAs containing inclusions at one or the other position or at both positions.

Another implication of the genomic sequence reported here is worthy of note. The sequence defines the positions of three introns (Figs. 3 and 4). One of these, the one involved in alternative splicing, occurs at the end of a type III repeat, while the other two both occur in the middle of homologous repeats. Therefore, the repeating structure of this region of fibronectin cannot be explained by the simplest model in which each type III repeat arises from a single exon.

We thank Norma Neff, Ihor Lemischka, and John Gardner for advice and useful discussions, Jennifer Lee for her excellent technical assistance, Richard Black for manuscript preparation, and Phil Sharp for critical reading of the manuscript. This work was supported by a grant from the National Cancer Institute (PO1 CA26712). J.E.S. was supported by a Damon Runyon Postdoctoral Fellowship.

1. Hynes, R. O. & Yamada, K. M. (1982) *J. Cell. Biol.* **95,** 369–377.
2. Yamada, K. M. (1983) *Annu. Rev. Biochem.* **52,** 761–799.
3. Tamkun, J. W. & Hynes, R. O. (1983) *J. Biol. Chem.* **258,** 4641–4647.
4. Atherton, B. T. & Hynes, R. O. (1981) *Cell* **25,** 133–141.
5. Hayashi, M. & Yamada, K. M. (1981) *J. Biol. Chem.* **256,** 11291–11300.
6. Schwarzbauer, J. E., Tamkun, J. W., Lemischka, I. R. & Hynes, R. O. (1983) *Cell* **35,** 421–431.
7. Kornblihtt, A. R., Vibe-Pedersen, K. & Baralle, F. E. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3218–3222.
8. Frischauf, A. M., Kehrach, H., Poustka, A. M. & Murray, N. (1983) *J. Mol. Biol.* **170,** 827–842.
9. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
10. Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L. & Treiber, C. (1970) *Virology* **40,** 734–744.
11. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3,** 2303–2308.
12. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15,** 687–701.
13. Seed, B., Parker, R. C. & Davidson, N. (1982) *Gene* **19,** 201–209.
14. Grosveld, F. G., Dahl, H. H. M., deBoer, E. & Flavell, R. A. (1981) *Gene* **13,** 227–237.
15. Vogelstein, B. & Gillespie, D. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 615–619.
16. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
17. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
18. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
19. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–580.
20. Hirano, H., Yamada, Y., Sullivan, M., DeCrombrugghe, B., Pastan, I. & Yamada, K. M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 46–50.
21. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–471.
22. Sharp, P. A. (1981) *Cell* **23,** 643–646.
23. Rozek, E. E. & Davidson, N. (1983) *Cell* **32,** 23–34.
24. King, G. R. & Piatigorsky, J. (1983) *Cell* **32,** 707–712.
25. Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wabl, M., Raschke, W. & Tonegawa, S. (1981) *Cell* **24,** 353–365.
26. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20,** 313–319.
27. Crabtree, G. R. & Kant, J. (1982) *Cell* **31,** 159–166.
28. Carlson, M. & Botstein, D. (1982) *Cell* **28,** 145–154.
29. Young, R. A., Hagenbuchle, O. & Schibler, U. (1981) *Cell* **23,** 451–458.
30. Nabeshima, Y., Fujii-Kuriyama, Y., Muramatsu, M. & Ogata, K. (1984) *Nature (London)* **308,** 333–338.
31. Ziff, E. B. (1980) *Nature (London)* **287,** 491–499.
32. Petersen, T. E., Thogersen, H. C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L. & Magnusson, S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 137–141.
33. Kornblihtt, A. R., Vibe-Pedersen, K. & Baralle, F. E. (1984) *EMBO J.* **3,** 221–226.