REFERENCES AND NOTES

1. W. M. Bayliss, *J. Physiol. (London)* **28**, 220 (1902).
2. G. Osol, R. Osol, W. Halpern, in *Essential Hypertension*, K. Aoki, Ed. (Springer-Verlag, Tokyo, 1986), pp. 107–113; D. R. Harder, R. Gilbert, J. H. Lombard, *Am. J. Physiol.* **253**, F778 (1987); J. M. Quayle, J. McCarron, W. Halpern, M. T. Nelson, *Biophys. J.* **57**, 301a (1990).
3. J. E. Brayden and G. C. Wellman, *J. Cereb. Blood Flow Metab.* **9**, 256 (1989).
4. M. T. Nelson, J. B. Patlak, J. F. Worley, N. B. Standen, *Am. J. Physiol.* **259**, C3 (1990).
5. J. N. Barrett, K. L. Magleby, B. S. Pallotta, *J. Physiol. (London)* **331**, 211 (1982); J. J. Singer and J. V. Walsh, *Pfluegers Arch.* **408**, 98 (1987); A. Carl and K. M. Sanders, *Am. J. Physiol.* **257**, C470 (1989); S. L. Hu, Y. Yamamoto, C. Y. Kao, *J. Gen. Physiol.* **94**, 849 (1989); R. Latorre, A. Oberhauser, P. Labarca, O. Alvarez, *Annu. Rev. Physiol.* **51**, 385 (1989).
6. U. Quast and N. S. Cook, *Trends Pharmacol. Sci.* **10**, 431 (1989).
7. R. Inoue, K. Kitamura, H. Kuriyama, *Pfluegers Arch.* **405**, 173 (1985); C. D. Benham, T. B. Bolton, R. J. Lang, T. Takewaki, *J. Physiol. (London)* **371**, 45 (1986); D. J. Beech and T. B. Bolton, *ibid.* **418**, 293 (1989).
8. Patch-clamp recordings were made with membrane patches excised in outside-out and inside-out configurations [O. P. Hamill, A. Marty, E. Neher, B. Sakmann, F. J. Sigworth, *Pfluegers Arch.* **391**, 85 (1981)] from enzymatically isolated smooth muscle cells of cerebral arteries at room temperature (*10, 24*).
9. Membrane depolarization from 0 to +50 mV increased the open state probability of these K_{Ca} channels by 67-fold (mean of three experiments) when intracellular Ca^{2+} was low (<100 nM). This voltage dependence is very similar to that reported for K_{Ca} channels in other tissues (*5, 7, 10*).
10. P. D. Langton, M. T. Nelson, Y. Huang, N. B. Standen, *Am. J. Physiol.* **260**, H927 (1991).
11. D. C. Ogden and D. Colquhoun, *Proc. R. Soc. London, Ser. B* **225**, 329 (1985); J. M. Quayle, N. B. Standen, P. R. Stanfield, *J. Physiol. (London)* **405**, 677 (1988).
12. C. E. Miller, E. Moczydlowski, R. Latorre, M. Phillips, *Nature* **313**, 316 (1985); A. L. Blatz and K. L. Magleby, *Trends Neurosci.* **10**, 463 (1987); R. J. Kovacs and M. T. Nelson, *Am. J. Physiol.* **261**, H604 (1991).
13. D. R. Harder, *Circ. Res.* **55**, 197 (1984).
14. Middle cerebral arteries were isolated, cannulated, and pressurized to 75 mmHg transmural pressure [B. R. Duling, R. W. Gore, R. G. Dacey, D. N. Damon, *Am. J. Physiol.* **241**, H108 (1981); G. Osol and W. Halpern, *ibid.* **249**, H914 (1985)]. Arterial diameter was measured with a video dimension analyzer. Membrane potential was recorded simultaneously with intracellular glass electrodes (*3*).
15. A. Galvez *et al.*, *J. Biol. Chem.* **265**, 11083 (1990).
16. A. L. Blatz and K. L. Magleby, *Nature* **323**, 718 (1986).
17. J. G. McCarron, J. M. Quayle, W. Halpern, M. T. Nelson, *Am. J. Physiol.* **261**, H287 (1991).
18. N. B. Standen *et al.*, *Science* **245**, 177 (1989).
19. D. R. Harder, *Circ. Res.* **60**, 102 (1987); G. M. Rubanyi, *Am. J. Physiol.* **255**, H783 (1988).
20. We disrupted the endothelium by placing an air bubble in the lumen for 1 min and then perfusing the lumen with distilled water for 30 s. Damage to the endothelium was verified by the absence of a dilator response to acetylcholine (1 μM) after myogenic tone developed.
21. Tone in isolated mesenteric and cerebral arteries is extremely sensitive to small changes in membrane potential [M. T. Nelson, N. B. Standen, J. E. Brayden, J. F. Worley III, *Nature* **336**, 382 (1988); J. E. Brayden, *Circulation* **84** (suppl. III), II-406 (1991) (*3, 4*). Thus, a change in membrane potential of only a few millivolts would be expected to cause a substantial change in diameter.
22. At intermediate levels of myogenic tone (pressure: 40 mmHg; diameter: 86 ± 4% of maximum; $n = 3$) the vasoconstrictor response to IBX (100 nM) was reduced (16 ± 6% decrease in diameter) compared to responses in arteries pressurized to 75 mmHg, which further supports a direct relation between the degree of myogenic tone and the extent of activation of K_{Ca} channels.
23. M. T. Nelson, Y. Huang, J. E. Brayden, J. Hescheler, N. B. Standen, *Nature* **344**, 770 (1990).
24. J. F. Worley, J. M. Quayle, N. B. Standen, M. T. Nelson, *Am. J. Physiol.* **261**, H1951 (1991).
25. P. C. Johnson, *Circ. Res.* **59**, 483 (1986).
26. G. A. Meininger, D. C. Zawieja, J. C. Falcone, M. A. Hill, J. P. Davey, *Am. J. Physiol.* **261**, H950 (1991).
27. A computer model developed by J. Patlak and M. T. Nelson of the ionic determinants of membrane potential in vascular smooth muscle provided insight that helped to develop this proposal.
28. In recent experiments, we found that coronary arteries isolated from rats develop substantial pressure-induced myogenic tone. In the presence of tone, CTX (50 nM) caused a significant decrease in vascular diameter (19 ± 4%; $n = 3$).
29. W. C. Cole, A. Carl, K. M. Sanders, *Am. J. Physiol.* **257**, C481 (1989); L. Toro, M. Amador, E. Stefani, *ibid.* **258**, H912 (1990); L. Toro and E. Stefani, *J. Bioenerg. Biomembr.* **23**, 561 (1991); H. Kume and M. I. Kotlikoff, *Am. J. Physiol.* **261**, C1204 (1991).

30. We thank A. Bonev, J. Asbury, and A. Zygmunt for technical assistance; J. Patlak, R. Gibbons, N. Standen, J. Hescheler, J. Daut, G. Osol, and M. Murphy for discussions and comments; Lederle Laboratories and American Cyanamid for CTX and IBX, A. Scriabine of Miles Laboratories for nimodipine, and Smith-Kline/Beecham Pharmaceutical Co. for cromakalim and lemakalim. Supported by grants from the NIH (HL 35911 and 44455), NSF (DCB-90195663), NATO, and an Established Investigator Award of the American Heart Association to J.E.B.

23 December 1991; accepted 3 March 1992

Isolation of a Complementary DNA That Encodes the Mammalian Splicing Factor SC35

Xiang-Dong Fu and Tom Maniatis

The mammalian splicing factor SC35 is required for the first step in the splicing reaction and for spliceosome assembly. The cloning and characterization of a complementary DNA encoding this protein revealed that it is a member of a family of splicing factors that includes mammalian SF2/ASF. This family of proteins is characterized by the presence of a ribonucleoprotein (RNP)-type RNA binding motif and a carboxyl-terminal serine-arginine–rich (SR) domain. A search of the DNA sequence database revealed that the thymus-specific exon (E_T) of the c-*myb* proto-oncogene is encoded on the antisense strand of the SC35 gene.

Splicing of nuclear pre-messenger RNA (pre-mRNA) takes place in complex ribonucleoprotein particles called spliceosomes, which contain pre-mRNA, the small nuclear ribonucleoproteins (snRNPs) U1, U2, U4/6, and U5, and non-snRNP proteins (*1*). Three non-snRNP mammalian splicing factors have been purified to homogeneity and characterized, namely SF2/ASF (*2*), U2AF (*3*) and SC35 (*4, 5*). The SC35 protein is required for formation of the earliest adenosine triphosphate (ATP)–dependent splicing complex (A complex) (*4*) and interacts with spliceosomal components bound to both the 5′ and 3′ splice sites during spliceosome assembly (*5*). In addition, SC35 is required for the ATP-dependent interactions of both U1 and U2 snRNPs with pre-mRNA.

We purified SC35 using standard chromatographic procedures (*6*) and generated tryptic peptides from the purified protein. These tryptic peptides were purified, and two were sequenced. The sequence of the first five NH_2-terminal amino acid residues of peptide 1 (Fig. 1) was identical to a sequence present in the splicing factor SF2/ASF (*2*), while the remaining 12 residues of the peptide were not present in this protein. The amino acid sequence of peptide 2 was identical to another SF2/ASF peptide (Fig. 1). Therefore, the purified SC35 preparation used to obtain the tryptic peptides was either contaminated with SF2/ASF, or the two proteins share common sequences.

We used the peptide 1 sequence to design oligonucleotide probes. One (probe A) corresponded to the SF2/ASF DNA sequence at the NH_2-terminus of peptide 1, and the other (probe B) corresponded to the unique region in the middle of peptide 1 (Fig. 1). We then used both probes simultaneously to screen a human primary cDNA library (*7, 8*). Of the five cDNA clones obtained, four hybridized to probe A and one to probe B. None of the clones hybridized to both probes. DNA sequence analysis of the two classes of cDNA clones revealed that the four clones that hybridized to probe A were SF2/ASF, while the sequence of the clone that hybridized to probe B (p1B) did not correspond to any known gene (Fig. 2A). The nucleotide sequence of the p1B cDNA clone encoded peptide 1, but did not contain the nucleotide sequence that corresponded to probe A or to the peptide 2 sequence found in SF2/ASF. Thus, the pooled SC35-containing MonoQ fractions were indeed contaminated by SF2/ASF, and

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138.

the p1B cDNA clone may therefore encode the splicing factor SC35.

The p1B cDNA clone has a single open reading frame (ORF) of 221 amino acids and there is no stop codon prior to the first AUG codon of the ORF (Fig. 2A). The sequence surrounding this AUG is an excellent match to the Kozak translation initiation consensus (9), and the amino acid sequence composition of the protein encoded by the open reading frame is very similar to that determined for purified SC35 (*10*). The deduced amino acid sequence contains a typical RNA binding RNP consensus sequence at the NH_2-terminus, a proline-glycine–rich sequence in the middle, and a large serine-arginine–rich domain at the COOH-terminus (Fig. 2, A and B). All three of these regions are conserved between SC35 and SF2/ASF (Fig. 2B). The *Drosophila* splicing regulator Tra-2 also contains the RNP consensus sequence and the serine-arginine–rich domain, but not the proline-glycine–rich region (*11*) (Fig. 2, B and C). The similarities between SC35 and SF2/ASF explain why the two proteins have similar chromatographic properties and suggest that p1B encodes a splicing factor related to SF2/ASF. Recently, SF2/ASF and SC35 were shown to be members of a family of splicing factors that are conserved between *Drosophila* and man (*12*). Five proteins from human cells were found to contain an RNP consensus sequence and a serine-arginine–rich region. One of these proteins (SRP33) appears to be identical to SC35 (*12*).

The SC35 cDNA is 1.9 kb in length (Fig. 2A), but we detected RNA species of 3.0 kb and 2.2 kb in all cell lines examined. A third transcript of 1.7 kb was also detected in some cell lines (Fig. 3). These RNAs may be alternatively processed forms of SC35 mRNA.

On the basis of the cDNA sequence, p1B encodes a protein of 221 amino acids (25.6 kD). This value is considerably lower than the molecular size of 35 kD determined by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). However, the protein generated by in vitro translation of recombinant SC35 RNA comigrated on gels with purified native SC35 (*13*). A similar discrepancy between the molecular size predicted from the cDNA and that determined by SDS-PAGE was reported for SF2/ASF (*2*). The recombinant SC35 protein containing a 12–amino acid Myc peptide tag was detected in transfected HeLa cells with an anti–Myc tag monoclonal antibody and co-localized with native SC35 to nuclear speckles (*4, 5, 13*). Thus, the molecular size and nuclear localization properties of the recombinant SC35 are indistinguishable from those of native SC35.

The cloned SC35 expressed in insect cells with a baculovirus vector (*14*) complemented SC35-depleted nuclear extracts for

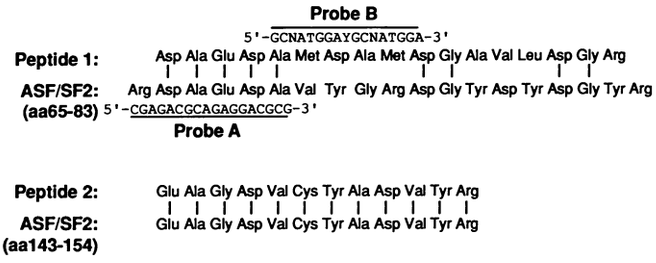

Fig. 1. Peptide sequences from a fast protein liquid chromatography (FPLC) MonoQ fraction that contained SC35. The peptide sequences were compared with those present in SF2/ASF (*2*). Synthetic deoxyoligonucleotide probes are shown aligned with the corresponding amino acid sequences. MonoQ fractions containing SC35 were combined, concentrated on a Centricon 30, resolved in a 12.5% SDS-PAGE, and blotted onto nitrocellulose. The proteins on the blot were stained with 0.1% Ponceau S (Sigma) in 1% acetic acid, and the strip that contained SC35 was cut out. The sample was digested with trypsin, and the peptides were isolated by high-performance liquid chromatography (HPLC) and sequenced in the Harvard Microchemistry Facility.

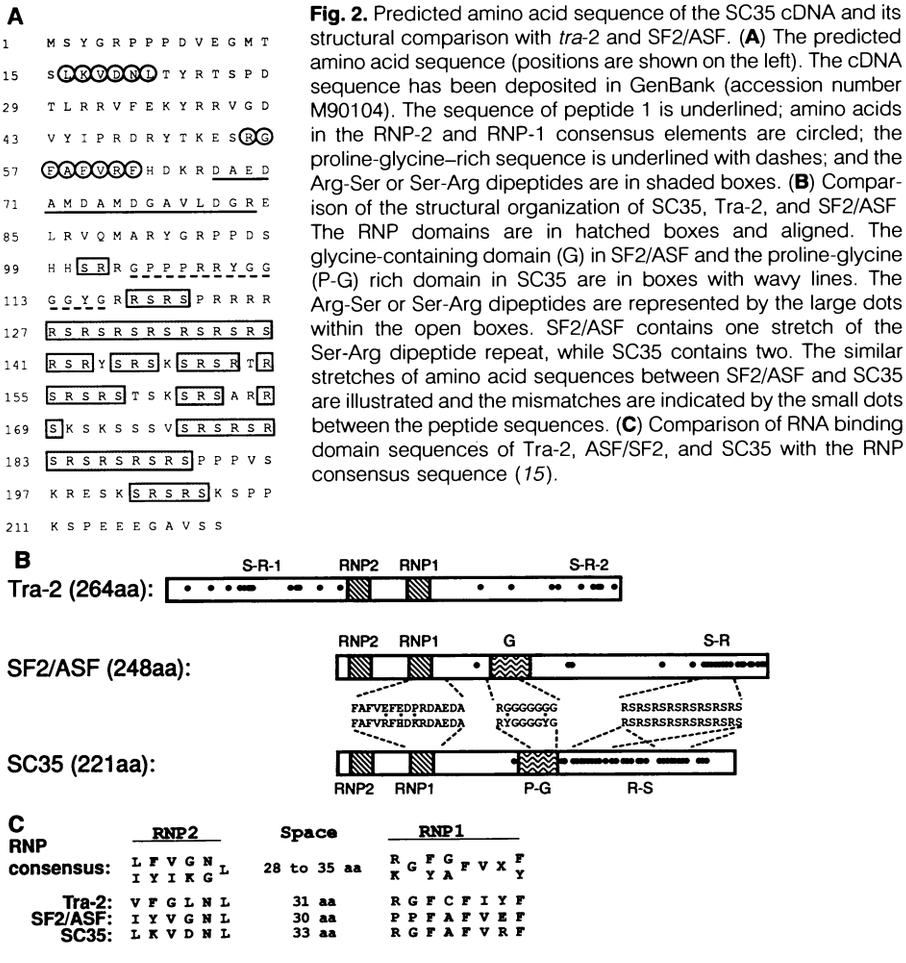

Fig. 2. Predicted amino acid sequence of the SC35 cDNA and its structural comparison with *tra*-2 and SF2/ASF. (**A**) The predicted amino acid sequence (positions are shown on the left). The cDNA sequence has been deposited in GenBank (accession number M90104). The sequence of peptide 1 is underlined; amino acids in the RNP-2 and RNP-1 consensus elements are circled; the proline-glycine–rich sequence is underlined with dashes; and the Arg-Ser or Ser-Arg dipeptides are in shaded boxes. (**B**) Comparison of the structural organization of SC35, Tra-2, and SF2/ASF The RNP domains are in hatched boxes and aligned. The glycine-containing domain (G) in SF2/ASF and the proline-glycine (P-G) rich domain in SC35 are in boxes with wavy lines. The Arg-Ser or Ser-Arg dipeptides are represented by the large dots within the open boxes. SF2/ASF contains one stretch of the Ser-Arg dipeptide repeat, while SC35 contains two. The similar stretches of amino acid sequences between SF2/ASF and SC35 are illustrated and the mismatches are indicated by the small dots between the peptide sequences. (**C**) Comparison of RNA binding domain sequences of Tra-2, ASF/SF2, and SC35 with the RNP consensus sequence (*15*).

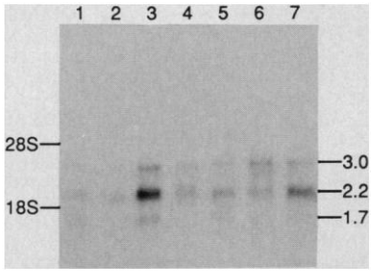

Fig. 3. Northern (RNA) blot analysis. Polyadenylated$^+$ RNA (2 μg) from different cell types was probed with antisense SC35 riboprobe. Lane 1, LCC-pk1 (pig kidney epithelial); lane 2, AtT20 (mouse pituitary tumor); lane 3, HT29 (human adenocarcinoma); lane 4, HG2 (human hepatocarcinoma); lane 5, Mel (human melanoma); lane 6, A10 (rat embryonic smooth muscle); lane 7, L6 (rat skeletal muscle myoblast). SC35 riboprobe was prepared by in vitro transcription of Eco RV–linearized pSP73-SC35 with T7 RNA polymerase (Promega) and was used to carry out hybridization as described (*8*).

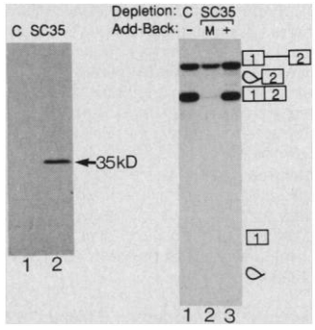

Fig. 4. Expression of SC35 and in vitro complementation of SC35-depleted nuclear extracts. Insect sf9 cells (10^4) infected with the baculovirus-expressing bicoid (C, control) or with the baculovirus-expressing SC35 were lysed in SDS buffer, and the soluble proteins were loaded onto a 12.5% SDS-polyacrylamide gel, transferred to nitrocellulose, and probed with anti-SC35 (left). For complementation analysis (right), HeLa cell nuclear extracts were depleted with a control monoclonal antibody that does not bind to SC35 (C, control depletion) or with anti-SC35. The SC35-depleted extracts were complemented with sf9 extracts that contained bicoid (M, mock) or SC35 (+). No SC35 (−) was added to the control antibody–depleted extract.

Fig. 5. Antisense sequence similarity between SC35 and the c-*myb* E_T exon. (**A**) Antisense sequence similarity between the human SC35 transcript and the chicken c-*myb* E_T exon (*18*). Relative positions of the nucleotides from the 5′ ends of corresponding cDNAs are indicated. Mismatches are indicated by stars. A portion of PCR-amplified cDNA from chicken (*19*) is shown and aligned with the similar E_T exon sequence. The chicken c-*myb* E_T exon is fused to the first exon (E1, lowercase) of the major c-*myb* transcript in chicken. (**B**) Antisense sequence similarity between the human SC35 transcript and a segment of E_T exon–containing human genomic DNA (*19*). A portion of amplified PCR cDNA from human (*21*) is shown and aligned with the similar E_T exon sequence. The human c-*myb* E_T exon is fused to the second exon (E2, lowercase) of the major c-*myb* transcript in human.

A

```
Chicken                  1                                              40
c-myb ET:        5'- GGCGGCCGTA GCGGGCCATC TGCACGCGGA GCTCGCGGCC
                                    *     *     *  *
Human 3'-CTCAGGCCCC CCGCCGGCAT CGCGCGGTAA ACGTGGGCGT CGAGCGCCGG
SC35:            438

         41                                                          90
         ATCCAGCACG GCCCCGTCCA TGGCGTCCAT CGCGTCCTCG GCGTCGCGTT
         *           *     *     *     *                 *
         CAGGTCGTGC CGGGGCAGGT ACCGTAGGTA TCGCAGGAGT CGCAGCGCGA

                                    PCR cDNA(chicken)
                                      ET              E1
                           5'-TGCGAACCCG CGGCTCctcg gcacttcgct- 3'
         91                                    125
         TGTCGTGGAA ACGGA-CGAA TGCGAACCCG CGGCTCTCCT TGGTGTAGCG- 3'
              *       *     *     *     **
         ACAGCAC-TT TCGCTTGCTT CCGCTTCGGC GCCCTGAGGA ACCACATCGC- 5'
                                                314
```

B

```
Human
genomic DNA              1                                              39
containing ET:   5'- GGCGGCCGTA GCGCGCCATT TGCAC-CGCA GCTCGCCGGC
                                                 *
Human 3'-CTCAGGCCCC CCGCCGGCAT CGCGCGGTAA ACGTGGGCGT CGAGCGCCGG
SC35:            438

                                    PCR cDNA(human)
                                      ET              E2
                           5'- TGGCATCCAT AGCagtgacg aggatgatga- 3'
         40                                    72
         GTCCAGCACG GCCCCGTCCA TGGCATCCAT AGCGTCCTCA GCGTCGCGCT

         CAGGTCGTGC CGGGGCAGGT ACCGTAGGTA TCGCAGGAGT CGCAGCGCGA


         90
         TGTCGTG-AA AGCGAACGAA GGCGAAGCCG CGGGACTCCT TGGTGTAGCG- 3'

         ACAGCAC-TT TCGCTTGCTT CCGCTTCGGC GCCCTGAGGA ACCACATCGC- 5'
                                                314
```

splicing. A recombinant baculovirus expressing the *Drosophila* bicoid protein was used as a control. Infected cells were lysed and the whole cell extracts analyzed by immunoblotting. A monoclonal antibody to SC35 (anti-SC35) detected a 35-kD protein in extracts from cells infected with the baculovirus containing SC35, but not in extracts from cells infected with the virus carrying the bicoid cDNA (Fig. 4).

Nuclear extracts were prepared from both bicoid- and SC35-expressing baculovirus-infected cells and tested for the SC35-complementing activity. The expressed recombinant SC35 was specifically recognized by anti-SC35 (*13*). HeLa cell nuclear extracts were depleted with either a control antibody or with anti-SC35 (Fig. 4). Depletion with the control antibody had no effect on splicing of a β-globin pre-mRNA, while depletion with anti-SC35 inhibited the splicing activity of the nuclear extracts (Fig. 4). Mock complementation of the SC35-depleted nuclear extracts with extracts from the bicoid-expressing baculovirus-infected cells did not rescue splicing. In contrast, addition of an extract that contained the recombinant SC35 protein restored the splicing activity (Fig. 4). Therefore, we conclude that the p1B cDNA encodes functional SC35.

The RNP motif in other RNA binding proteins is essential for RNA binding (*15*). Both SC35 and SF2/ASF bind to RNA, but do not appear to recognize specific sequences in naked pre-mRNA as assayed by ultraviolet cross-linking (*2*, *13*). However, these factors may interact with specific snRNA or pre-mRNA sequences in the context of spliceosome. Although the function of the serine-arginine–rich sequence is not known, this structural motif has been implicated in the subnuclear localization of *Drosophila* splicing regulators Tra and Su(W^a) (*16*). Consistent with this function, a long stretch of serine-arginine–rich sequence is present in SC35, and the protein localizes to speckles (*4*). Other members of this serine-arginine–rich family display the same speckled distribution in the nucleus (*17*).

We searched DNA sequence databases for sequences similar to SC35 and found that a 125-nt stretch of SC35 is strikingly similar to the antisense DNA strand of the chicken c-*myb* E_T exon (*18*) (Fig. 5A). The human genomic DNA fragment containing the c-*myb* E_T region has been cloned (*19*) and the sequence of the transcribed DNA strand is identical, except for one mismatch, to antisense SC35 over the region reported (Fig. 5B). Thus, the c-*myb* E_T exon directly overlaps the SC35 sequence, indicating that both strands of this genomic sequence are transcribed. This coding potential of the antisense strand of the E_T exon has been noted, and cDNA clones of the divergently transcribed RNAs have been isolated from chicken and human cells with an E_T exon–specific probe (*20*). Indeed, the sequence of the SC35 cDNA clone reported here is identical (except for a single-base polymorphism) to the human cDNA containing the antisense E_T exon (*20*).

The c-*myb* E_T exon and the other c-*myb* exons are located on different chromosomes in both chicken and human cells (*19*). Despite this discontinuity in the coding sequences, an RNA containing both the E_T exon and the remaining 15 c-*myb* exons has been detected in normal chicken thymus (*18*, *19*) (Fig. 5A). The results of polymerase chain reaction (PCR) experiments suggest that a similar hybrid RNA may also be produced in human thymus cells (*21*) (Fig. 5B), but definitive proof for this conclusion has not been obtained. If hybrid E_T exon–c-*myb* mRNAs do exist, they must be the result of either a cell-specific DNA translocation event or of trans-splicing as suggested previously (*19*). The observation that a potentially trans-spliced exon of the c-*myb* proto-oncogene is encoded by the antisense strand of a gene encoding an essential splicing factor may be of functional significance.

REFERENCES AND NOTES

1. M. R. Green, *Annu. Rev. Cell Biol.* **7**, 559 (1991); R. Reed, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8031 (1990); S. W. Ruby and J. Abelson, *Trends Genet.* **7**, 79 (1991); U. Vijayraghavan *et al.*, *Genes Dev.* **3**, 1206 (1989).
2. A. R. Krainer *et al.*, *Genes Dev.* **4**, 1158 (1990); *Cell* **62**, 35 (1990); *ibid.* **66**, 383 (1991); H. Ge and J. L. Manley, *ibid.* **62**, 25 (1990); H. Ge *et al.*, *ibid.*

66, 373 (1991).
3. P. D. Zamore and M. R. Green, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9243 (1989); *EMBO J.* **10**, 207 (1991).
4. X.-D. Fu and T. Maniatis, *Nature* **343**, 437 (1990); D. L. Spector *et al.*, *EMBO J.* **10**, 3467 (1991).
5. X.-D. Fu and T. Maniatis, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1725 (1992).
6. Purification of SC35. HeLa nuclear extract (10 ml) was mixed with an equal volume of saturated ammonium sulfate (AS), and the mixture was stirred at 4°C for 1 hour. The precipitate was removed by centrifugation at 4000 rpm ($4657g_{max}$) for 30 min. in a Beckman centrifuge H-6000A. When assayed by immunoblotting, SC35 was found in the soluble fraction. The SC35-containing fraction was applied to a 3-ml (packed volume) phenyl-Sepharose (Pharmacia) column equilibrated in buffer A (20 mM tris-HCl, pH 7.9; 100 mM KCl; 0.2 mM EDTA) plus 1.7 M AS. After washing the column with 5 ml of the same buffer, proteins were step-eluted with buffer A plus 0.2 M incremental decreases in AS concentration or with buffer A alone. SC35 was eluted in two peaks at 1.3 M and 0.5 M AS. The 1.5 to 1.1 M fractions containing SC35 were combined, dialyzed overnight against buffer B (buffer A plus 5% glycerol), concentrated on a Centriprep 30 (Amicon), and loaded onto a MonoQ column (Pharmacia) equilibrated in the same buffer. The MonoQ column was developed with a linear KCl concentration gradient from 0.1 to 1.0 M KCl in buffer B, and 1-ml fractions were collected. SC35 was eluted in one peak at 0.35 M KCl, and the pooled MonoQ fractions were used for peptide sequencing analysis in this study. The peak fraction was further fractionated on a glycerol gradient, concentrated on a Centricon 30, and loaded onto a 10 to 30% linear glycerol gradient, followed by centrifugation at 47,000 rpm for 23 hours at 4°C in a Beckman SW55 rotor. After centrifugation, 0.5-ml fractions were collected. SC35 was apparently homogeneous in one of the fractions and was used for protein composition determination or for functional analyses.
7. Isolation of SC35 cDNA clones. Two DNA oligomers (probe A and probe B) were synthesized according to the sequence of peptide 1 (Fig. 1). These two probes were end-labeled with [^{32}P]ATP and T4 kinase. The labeled probes were used to screen 0.5 × 10^6 plaques from a human B cell primary cDNA library with the TMACl (tetramethylammonium chloride) hybridization protocol (*8*). The SC35 insert was isolated from phage DNA, subcloned into the Eco RI site of the pSP73 vector (pSP73-SC35), and sequenced.
8. J. Sambrook, E. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ed. 2, 1989).
9. M. Kozak, *Cell* **44**, 283 (1986).
10. The open reading frame of p1B cDNA contains 24% arginine and 20% serine residues, compared to 27% arginine and 18% serine of purified SC35. The calculated isoelectric point for the protein encoded by the clone is 12.38.
11. H. Amrein, *Cell* **55**, 1025 (1989); T. J. Goralski *et al.*, *ibid.* **56**, 1011 (1989).
12. A. Mayeda, A. M. Zahler, A. R. Krainer, M. B. Roth, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1301 (1992); A. Zahler, W. Lane, M. Roth, *Genes Dev.*, in press.
13. X.-D. Fu and T. Maniatis, unpublished data.
14. M. D. Summers and G. E. Smith, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experiment Station, Bulletin no. 1555 (1987). An Eco RV–Pvu II fragment that contained SC35 from pSP73-SC35 was cloned into the Nhe I site of the baculovirus expression vector pJVP10Z. The translation initiation signal of SC35 was used in this construct. Recombinant virus expressing SC35 was produced according to the manual. Insect cell nuclear extracts containing SC35 or bicoid were prepared in the same way as HeLa nuclear extracts, and 5 ml of nuclear extract were obtained from 100 ml of infected sf9 cells (1 × 10^8). Depletion and complementation analyses were as described (*4, 5*) and modified (*5*).
15. G. Dreyfus *et al.*, *Trends Biochem. Sci.* **13**, 86 (1988); D. J. Kenan *et al. ibid.* **16**, 214 (1991); C. C. Query *et al.*, *Cell* **57**, 89 (1989).
16. H. Li and P. M. Bingham, *Cell* **67**, 335 (1991).
17. M. B. Roth *et al.*, *J. Cell Biol.* **111**, 2217 (1990); *ibid.* **115**, 587 (1991).
18. D. Rosson and E. P. Reddy, *Nature* **319**, 604 (1986).
19. M. Vellard *et al.*, *Oncogene* **6**, 505 (1991).
20. B. Perbal and M. Vellard, *C. R. Acad. Sci. Paris* **311**, 467 (1990); M. Vellard *et al.*, *ibid.* **313**, 591 (1991).
21. A. Sureau and B. Perbal, *ibid.* **312**, 323 (1991).
22. We thank A. Krainer, J. Manley, H. Ge, and M. Roth for communicating their unpublished results; W. Lane for amino acid composition and peptide sequencing; X.-F. Wang, H. Wu, and M.-J. Jin for advice on Northern blotting analyses and peptide tagging; A. Leza for computer sequence homology search; G. Shen-Ong for many insights into the c-*myb* proto-oncogene; R. Kriz for the human B cell primary cDNA library; K. Diepold for the recombinant virus–expressing bicoid; and J. Bruzik, J. Wu, M. Tian, H. Amrein, M-L. Hedley, T. Schaal, and members of the Maniatis laboratory for useful discussions and critical comments on the manuscript. Supported by NIH grant GM42231 to T.M.

15 January 1992; accepted 6 March 1992

Regulatory Elements That Control the Lineage-Specific Expression of *myo*D

David J. Goldhamer, Alexander Faerman, Moshe Shani, Charles P. Emerson, Jr.*

The molecular basis of skeletal muscle lineage determination was investigated by analyzing DNA control elements that regulate the myogenic determination gene *myo*D. A distal enhancer was identified that positively regulates expression of the human *myo*D gene. The *myo*D enhancer and promoter were active in myogenic and several nonmyogenic cell lines. In transgenic mouse embryos, however, the *myo*D enhancer and promoter together directed expression of a *lac*Z transgene specifically to the skeletal muscle lineage. These data suggest that during development *myo*D is regulated by mechanisms that restrict accessibility of *myo*D control elements to positive trans-acting factors.

The formation of differentiated tissue types during development involves the determination of specific cell lineages from multipotential progenitor cells followed by terminal differentiation, resulting in cell type–specific gene expression and function. The molecular mechanisms that regulate cell lineage determination are poorly understood. However, the mouse cell line C3H10T1/2 (10T1/2) has provided a model cell culture system (*1*) that has allowed identification of four mammalian genes [*myo*D (*2*), *myogenin* (*3*), *Myf-5* (*4*), and *MRF-4* (*5*)] that regulate the determination of the skeletal muscle lineage (*6*). These genes encode transcription factors that comprise a subgroup within the helix-loop-helix (HLH) superfamily of Myc-related DNA binding proteins (*7*). As transfected complementary DNAs (cDNAs), these factors induce myogenic conversion of multipotential 10T1/2 cells to stably determined populations of proliferative myogenic cells (*2–5*). Consistent with a function in determination, these myogenic regulatory genes are expressed exclusively in skeletal muscle lineages of the embryo, beginning at early stages of somite formation (*8–10*). Although the function of the proteins encoded by these genes is beginning to be elucidated (*6*), the transcriptional regulatory mechanisms that activate their expression in the skeletal muscle lineage of the embryo are unknown.

We have analyzed the cis-acting DNA sequences that regulate expression of the human *myo*D gene. A pWE15 cosmid library was screened at moderate stringency with a full-length mouse MyoD cDNA (*11*). This screen yielded four recombinants representing three unique overlapping clones that spanned a total of 40 kb. Sequence comparison with the human MyoD cDNA (*12*) identified the hybridizing species as *myo*D. The organization of the cosmid clone used in subsequent analyses (chMD-*13*) included ~25.5 kb of DNA upstream and 4 kb downstream of the *myo*D gene (Fig. 1A).

Transcriptional activity of the *myo*D promoter was assayed in 23A2 myoblasts, myogenic cells derived from the multipotential 10T1/2 cell line by 5-azacytidine treatment (*1*). We fused 2.5 kb of 5′ flanking sequence of *myo*D to the chloramphenicol acetyltransferase (CAT) reporter gene (−2.5CAT; Fig. 1A) (*13*), and CAT activity was assayed after transient transfection into proliferative 23A2 myoblasts (*14*). The *myo*D promoter is only weakly

D. J. Goldhamer and C. P. Emerson, Jr., Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, PA 19111.
A. Faerman and M. Shani, Institute of Animal Science, Volcani Center, P.O. Box 6, Bet Dagan, 50250 Israel.
*To whom correspondence should be addressed.