MOLECULAR AND CELLULAR BIOLOGY, Aug. 1990, p. 4068-4079
0270-7306/90/084068-12$02.00/0

Vol. 10, No. 8

An Alternative Non-Tyrosine Protein Kinase Product of the c-*src* Gene in Chicken Skeletal Muscle

THAMBI DORAI AND LU-HAI WANG*

Department of Microbiology, Mount Sinai School of Medicine, New York, New York 10029-6574

Received 25 October 1989/Accepted 9 May 1990

While the c-*src* locus is expressed as a 4.0-kilobase (kb) mRNA coding for pp60$^{c\text{-}src}$ in various chicken tissues, including embryonic muscle, it is expressed as a novel 3.0-kb mRNA in adult skeletal muscle. We have analyzed the primary structure of this alternatively transcribed and spliced c-*src* mRNA. The sequence revealed three open reading frames, with the previously defined c-*src* exons 1 through 5 or 6 comprising the third, on the 3′ untranslated region of this 3-kb mRNA. The exons coding for the tyrosine kinase domain of pp60$^{c\text{-}src}$ were excluded. On the 5′ side, 2 kb of sequence upstream from the previously defined exon 1 of the c-*src* gene was included in this mRNA. The start site for the 3-kb mRNA probably lies downstream of that for the 4-kb mRNA. The first reading frame of the 3.0-kb mRNA, called *sur* (for *src* upstream region), encoded a 24-kilodalton (kDa) protein product rich in cysteine and proline residues. In vitro analysis indicated that the 24-kDa *sur* protein was membrane associated. Antibodies to *sur* protein detected in vivo a 24-kDa muscle-specific protein which was developmentally regulated and corresponded to the switch from the 4-kb to the 3-kb c-*src* mRNA. A striking kinetic pattern of appearance of *sur* protein and disappearance of pp60$^{c\text{-}src}$ suggests that the expression of these two proteins is inversely related.

The proto-oncogene c-*src* is the cellular homolog of the oncogene v-*src* of Rous sarcoma virus and represents the prototype of a big family of membrane-associated tyrosine protein kinases (TPKs) in normal cells (reviewed in references 4, 22, and 26). The ubiquitous presence of c-*src* in all metazoa ranging from freshwater sponges to humans strongly implies that it performs an essential yet undefined role in developmental or normal cellular function (3, 41). The protein products of both c-*src* and v-*src* possess TPK activity and are phosphoproteins attached to the inner side of the plasma membrane via N-terminal myristic acid (11, 14, 28, 37, 60).

Despite their similarities in biochemical properties, the c-*src* protein product is fundamentally different from that of v-*src* in its transforming potential (29, 30, 57). The difference has been shown to stem from minor structural changes in the v-*src* protein (30, 42, 66). These structural variations in v-*src* correlate with its enhanced TPK activity and transforming ability (12, 15, 25, 34). The TPK activity has been shown to be absolutely required for the transforming function of v-*src* (32, 61) and is undoubtedly also essential for the normal function of c-*src*, although this has not been formally proved.

Several lines of evidence indicate that expression of c-*src* is developmentally regulated and may play an important role in differentiation of certain tissues. Studies of c-*src* expression at both RNA and protein levels show that it is widely expressed at a very low level in various tissues and is significantly higher in neural tissues, particularly in embryonic neurons (6, 23, 63, 70). Quantitative RNA analyses showed that the level of c-*src* mRNA is elevated in chicken embryonic brain at various stages of development (20, 50). With immunocytochemistry, pp60$^{c\text{-}src}$ was shown to localize in the process-rich molecular layer of the cerebellum of chicken embryos at the time of neuronal differentiation (18). Besides embryonic brain tissues of chicken, rat, and human (9, 31, 41, 47) high levels of pp60$^{c\text{-}src}$ were also observed in developing chicken neural retina (62). The appearance of pp60$^{c\text{-}src}$ at the onset of neuronal differentiation implies that it may be involved in maturation rather than proliferation of neuronal precursors.

In *Drosophila melanogaster*, there is also a specific accumulation of c-*src* mRNA in neural tissues and in smooth muscle cells after the cessation of cell proliferation (58). More recent studies showed that the fully differentiated postmitotic neurons express high levels of a structurally distinct and activated form of pp60$^{c\text{-}src}$ (6–8). The specific kinase activity of the pp60$^{c\text{-}src}$ from these neurons was 6 to 12 times higher than that derived from astrocytes. This unique form of pp60$^{c\text{-}src}$ was also found when cultured embryonal carcinoma cells were induced to elaborate neuronlike processes by treatment with retinoic acid (45).

Analyses of the c-*src* cDNA clones from mouse (48) and chicken (43) embryonic brain mRNAs revealed that this novel pp60$^{c\text{-}src}$ is synthesized from an alternatively spliced transcript with six additional codons between exons 3 and 4 of regular pp60$^{c\text{-}src}$. Other nonproliferating cells, such as platelets (21) and myeloid cells (19), also contain elevated pp60$^{c\text{-}src}$ TPK compared with other nonneuronal tissues, indicating again that pp60$^{c\text{-}src}$ fulfills an as yet undefined normal function other than mitogenicity in those fully differentiated cells.

In order to elucidate the normal function of pp60$^{c\text{-}src}$, we have initiated a study of the structure of c-*src* mRNA from various chicken tissues (69). The c-*src* mRNA is 4.0 kilobases (kb) in size, with a noncoding sequence of about 2.0 kb located at the 3′ end. While this 4.0-kb mRNA is expressed in various chicken tissues of both embryonic and adult origins, a novel 2.8- to 3.2-kb (hereafter referred to as 3.0-kb) c-*src* mRNA was found to be expressed specifically in chicken skeletal muscle. Results of our previous study indicated that (i) both the 4- and 3-kb c-*src* mRNAs are transcribed in the same polarity from the c-*src* locus, (ii) expression of the two mRNAs species is mutually exclusive, and the irreversible transition of expression from the 4- to the 3-kb mRNA occurs shortly before hatching of chicken,

* Corresponding author.

and (iii) the 3-kb mRNA lacks the region coding for the catalytic domain of the $pp60^{c\text{-}src}$ while it contains sequences derived from regions upstream from exon 1 of the 4-kb mRNA.

Based on these observations, we postulated that these two RNAs were generated under tissue-specific and developmental control by alternative splicing. We describe in the present study the isolation and sequencing of the cDNA of the 3-kb c-*src* mRNA and an initial characterization of its encoded product.

MATERIALS AND METHODS

Isolation of genomic c-*src* DNA fragments and Southern hybridization. The various fragments of the c-*src* DNA shown in Fig. 2 were isolated from appropriate restriction enzyme digests of molecular clones described previously (65, 69) under conditions recommended by the manufacturer. ^{32}P-labeled probes were prepared either by nick translation or by random oligonucleotide-primed synthesis (Promega-Biotec) of gel-purified DNAs. Southern hybridization was carried out as described previously (69).

Isolation of cDNA clones of c-*src* from muscle. Polyadenylated RNAs were isolated from pooled 14-day-old chicken pectoral muscle by the method described previously (68). Poly(A)-containing RNAs were fractionated in a sucrose gradient, and sample RNAs from each fraction were analyzed by the Northern (RNA blot) procedure for the presence of the 3-kb c-*src* mRNA. RNAs ranging from 18 to 30S, including peak fractions for the 3-kb RNA, were pooled and used as templates in cDNA library construction with oligo$(dT)_{12-18}$ as primers and λgt10 as the vector by published methods (71). Initially, cDNA clones were isolated by using c-*src* probe 5 (see Fig. 2) as the molecular probe for screening. Other overlapping clones were isolated from the same library by rescreening it with cDNA clone 3 or 2.

DNA sequencing. Recombinant bacteriophage DNAs were purified from plate lysates by using HflC-600 *Escherichia coli* cells as the host. The cDNA inserts were excised by using *Eco*RI and ligated into M13mp18 or M13mp19. Subclones in both orientations were identified, and overlapping deletion clones were prepared by the method of Dale et al. (16). These overlapping deletion clones were sequenced in both orientations to construct the entire sequence. Various modifications of sequencing methods, such as the use of dITP, 7-deaza-2′-deoxyguanosine-5′-triphosphate, and Sequenase (U.S. Biochemical Corp.), were introduced to resolve problem areas according to standard sequencing procedures offered by the manufacturers.

Construction of full-length c-*src* cDNA from muscle. A 1.3-kb 3′ *Bgl*II-*Eco*RI fragment from cDNA clone 3 (Fig. 1) was ligated to the 5′ *Eco*RI-*Bgl*II fragment of clone 2 and cloned into pGEM3 vector to form pGEM c-src^{m} (see Fig. 6A). The orientation of the insert was checked by restriction enzyme analysis.

Cell-free transcription of translation. Capped RNA transcripts were made from the *Hin*dIII-linearized pGEM c-src^{m} by using T7 RNA polymerase (Promega-Biotec). After digesting the template DNA with RNase-free RQI DNase (Promega Biotec), 500 ng of the transcript was incubated in 25 μl of a solution containing 20 μM amino acid mixture (minus methionine), 18 μl of nuclease-treated rabbit reticulocyte lysate, and 25 μCi of [^{35}S]methionine (1,200 Ci/mmol, 10 μCi/μl) at 30°C for 1 h. For processing experiments, 1 μl of stripped microsomal membranes of canine pancreas (2 processing equivalents; Promega Biotec) was included during in vitro translation. The reaction was stopped by transferring the tubes to 0°C, and samples of the reaction mixture were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Fractionation of translation products on sucrose step gradients and protease protection analysis. Fractionation and protection analyses were done essentially by the method of Addison (1). Protease protection experiments in the presence of canine pancreas microsomal membranes were done by the procedure of Sugano et al. (64).

Construction of the bacterial expression vector for the *sur* polypeptide. A 787-base-pair (bp) *Ban*I-*Ban*I fragment from cDNA clone 2 containing the 3′ 168 amino acids of the *sur* coding region was blunted by T4 DNA polymerase and cloned into *Cla*I-digested and blunted expression vector pSJH57 (see Fig. 6B), which contains the thermoinducible λ p_L promoter, by using *E. coli* CD646 cells as the host. In this construction, the sequence encoding the N-terminal 12 amino acids of lambda *c*II protein is fused to the *sur* sequence (10, 39). DNAs of potential positive clones were analyzed by hybridization and restriction mapping. Clones with the correct orientation for expression were selected and used to transform *E. coli* N4830, which was used as the expression host. Expression protein products were analyzed after heat induction and lysis of the bacteria, followed by SDS-PAGE. For large-scale preparation, bacteria were lysed and extracted with detergents (Nonidet P-40 and sodium deoxycholate) and salt, and the *sur* product remaining as insoluble precipitates was extracted by boiling in SDS-PAGE sample buffer and purified by SDS-PAGE according to published procedures (10, 39).

Immunization of rabbits was done either by injecting the gel-purified *sur* protein into lymph nodes or by injecting subcutaneously at multiple sites the crushed polyacrylamide gel pieces containing the *sur* protein. In either case, the protein preparation was mixed with complete freund adjuvant by standard protocols. Both methods yielded antisera with satisfactory titers and specificity.

Purification of immunospecific antibodies. The bacterially expressed *sur* protein was purified by preparative PAGE and blotted onto nitrocellulose paper electrophoretically. In an independent analysis, the *sur* antigen electroeluted from polyacrylamide gels was extensively dialyzed in phosphate-buffered saline to remove SDS and coupled to cyanogen bromide-activated Sepharose CL-4B to give a matrix containing about 5.0 μg of *sur* protein per ml of gel. Adsorption of specific antibodies in the anti-*sur* antiserum and elution of those antibodies were performed as described by Sambrook et al. (53).

Immunoprecipitation analysis. In vitro-translated and [^{35}S] methionine-labeled products were immunoprecipitated with polyclonal or affinity-purified anti-*sur* antiserum in RIPA buffer (50 mM Tris hydrochloride [Tris-HCl, pH 7.4], 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with Trasylol (100 kIU/ml), 1 mM phenylmethylsulfonyl fluoride, and 10 μg each of soybean trypsin inhibitor, leupeptin, and antipain per ml. Immunoprecipitation was also performed in parallel with *src* monoclonal antibodies 327 and GD11. Following an overnight incubation at 4°C with gentle shaking, the immune complexes were collected with protein A immobilized on Sepharose. For monoclonal antibodies, a secondary antibody (rabbit anti-mouse immunoglobulin G [IgG]) was used after primary antibody incubation as described by Lipsich et al. (44). The immune complexes adsorbed to protein A-Sepha-

rose were washed five times with RIPA buffer and analyzed by 12% SDS–PAGE.

For treatment with alkaline phosphatase, the immunoprecipitates were washed with assay buffer (50 mM Tris-HCl [pH 9.0], 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 1 mM spermidine) and incubated in the same buffer with 100 U of calf intestinal alkaline phosphatase at 37°C for 1 h. After incubation, the immunoprecipitates were washed four times with RIPA buffer and analyzed by SDS-PAGE. The immunoprecipitates were also treated with potato acid phosphatase as described by Cooper and King (13).

For RNase treatment of immunoprecipitates, incubation was done in a final volume of 100 μl at 37°C for 30 min in the presence of 10 mM Tris-HCl (pH 7.4) with a mixture of RNase A (40 μg; Boehringer Mannheim), RNase T_1 (2,000 U; Bethesda Research Laboratories [BRL]), and RNase T_2 (150 U; BRL). Finally, the precipitates were washed with RIPA buffer and processed for SDS-PAGE as described above.

Western immunoblotting. Various embryonic and young adult (2-week-old) chicken tissues were minced on ice in the presence of protease inhibitors, which included 1 mM phenylmethylsulfonyl fluoride, 100 kIU of aprotinin (Trasylol) per ml, and 10 μg/ml each of soybean trypsin inhibitor, leupeptin, and antipain. The minced tissues were boiled for 10 min in SDS-PAGE sample buffer, and the insoluble debris was removed by centrifugation. Concentrations of proteins were measured with a Bio-Rad Laboratories protein assay kit after the removal of interfering substances such as SDS and β-mercaptoethanol by ethanol precipitation of proteins.

After gel electrophoresis in 12% polyacrylamide gels, the separated proteins were transferred to nitrocellulose (0.45 μm pore size; Schleicher & Schuell) with a Trans-Blot apparatus (Bio-Rad) for 12 h at 40 V in a buffer containing 25 mM Tris-glycine (pH 8.3) and 20% methanol. Additional protein-binding sites on the nitrocellulose blot were blocked by 0.5% gelatin and 0.05% Tween 20 in Tris-buffered saline. The blots were then incubated with either preimmune serum or anti-*sur* or anti-*src* antiserum at a 1:250 dilution in blocking buffer for 2 h at room temperature. The blots were extensively washed with Tris-buffered saline containing 0.05% Tween 20 and treated with ^{125}I-protein A (ICN) as described by Wang (67). When monoclonal antibody 327 was used, the blots were treated with 1% nonfat dry milk and rabbit anti-mouse IgG (at a 1:1,000 dilution) and finally with ^{125}I-labeled protein A as described above.

RESULTS

Isolation of the cDNA clones of the c-*src* mRNAs in muscle. Our previous study showed that the 3.0-kb c-*src* mRNA is present exclusively in chicken skeletal muscle (69). This RNA was detected by using a v-*src* probe spanning exons 2 and 3. When the sequence content of the novel 3.0-kb mRNA was analyzed with various c-*src* genomic DNA probes, it was found that a 2.5-kb *Bgl*II-*Hin*dIII fragment, denoted the c-*src* probe 5 (see Fig. 2), hybridized strongly with the new mRNA, while the same probe hybridized only weakly with the commonly observed 4-kb c-*src* mRNA. This observation suggested that c-*src* exon 1 and/or considerable amounts of its upstream sequences are probably present in the 3.0-kb mRNA.

To elucidate the nature of this novel 3-kb mRNA, we constructed a cDNA library of the poly(A)-containing RNAs from 14-day-old chicken pectoral muscle. By using c-*src* probe 5 to screen the library, several cDNA clones were initially isolated. Two clones selected for further characterization are shown in Fig. 1 (clones 3 and 4). Rescreening the cDNA library with clone 3 as the probe yielded a 5′-overlapping clone of 1.5 kb (denoted clone 2 in Fig. 1). Further screening of the library with clone 2 yielded a small upstream 5′-overlapping clone of 0.5 kb (clone 1 in Fig. 1).

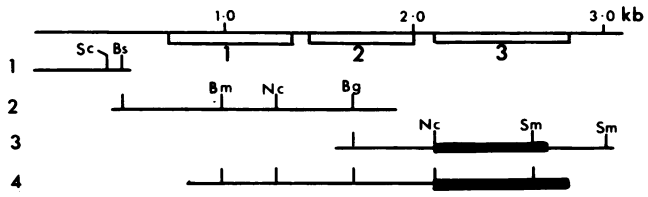

FIG. 1. Organization of the chicken muscle c-*src* cDNA clones. The top line shows the scale for the 3-kb c-*src* mRNA. The various overlapping clones are depicted with a brief restriction map. The solid boxes in clone 3 and clone 4 indicate c-*src* coding exons 2 through 5 and exons 2 through 6, respectively. The restriction sites are abbreviated as follows: Sc, *Sac*I, Bs, *Bst*NI; Bm, *Bam*HI; Nc, *Nco*I; Bg, *Bgl*II; Sm, *Sma*I. The three open reading frames found in the sequence are depicted by open boxes numbered 1, 2, and 3.

Figure 1 depicts the alignment of the overlapping cDNA clones derived by nucleotide sequencing. The arrangement also shows that the entire cDNA sequence derived from the overlapping clones could account for the size of the 3.0-kb RNAs detected in the Northern analyses. Sequencing analysis revealed that none of the four clones had the poly(A) tracts representing the 3′ ends of the mRNAs, most likely resulting from internal initiations during cDNA synthesis. Sequencing also revealed that $pp60^{c\text{-}src}$ exons 1 through 5 and 6 were present in clones 3 and 4, respectively (solid boxes in Fig. 1). In both clones, the $pp60^{c\text{-}src}$ exons were located on the 3′ side of the mRNAs.

To check the authenticity of these clones, each was used as a probe in Northern analysis of muscle RNA and found to detect only the previously described 2.8- to 3.2-kb muscle-specific mRNAs (data not shown). By using clone 3 cDNA as a probe in a Southern analysis of chicken genomic DNA, only the expected DNA fragments derived from the c-*src* locus were detected (see Fig. 4 in reference 69).

To further map the origins of the sequences in these cDNAs, they were used as probes to hybridize with various c-*src* genomic DNA fragments (Fig. 2) in Southern analyses. The entire c-*src* locus was divided into 17 fragments, as shown in the figure, and each cDNA clone was used to probe the Southern blots containing a panel of these fragments. The results are summarized in Fig. 2. The 5′ diverged sequences of the cDNA clones (sequences upstream from the solid boxes in Fig. 1) were shown to be derived from c-*src* probe 4 and 5 regions. The 3′ diverged sequence from clone 3 (sequences downstream from the solid box) hybridized with the 1.3-kb *Sac*I-*Mlu*I fragment (Fig. 2), implying that it was derived from either intron 5 or intron 6. Neither the 5′ nor the 3′ diverged sequences hybridized with the 4-kb c-*src* mRNAs from chicken embryo fibroblasts (CEF) or chicken brain (data not shown). The authenticity of the 5′ and 3′ diverged sequence in clone 3 was further confirmed by its hybridization only to the 2.8- to 3.0-kb mRNAs in Northern analysis (data not shown). These results show that the 3-kb mRNA as represented by these cDNA clones is generated by alternative RNA splicing from the normal c-*src* locus. It shares sequences with $pp60^{c\text{-}src}$ 4.0-kb mRNA exons 1 to 5 (clone 3) or 1 to 6 (clone 4) and differs from the 4.0-kb mRNA in its 5′ and 3′ sequences. The 5′ diverged sequence is denoted *sur*, standing for *src* upstream region.

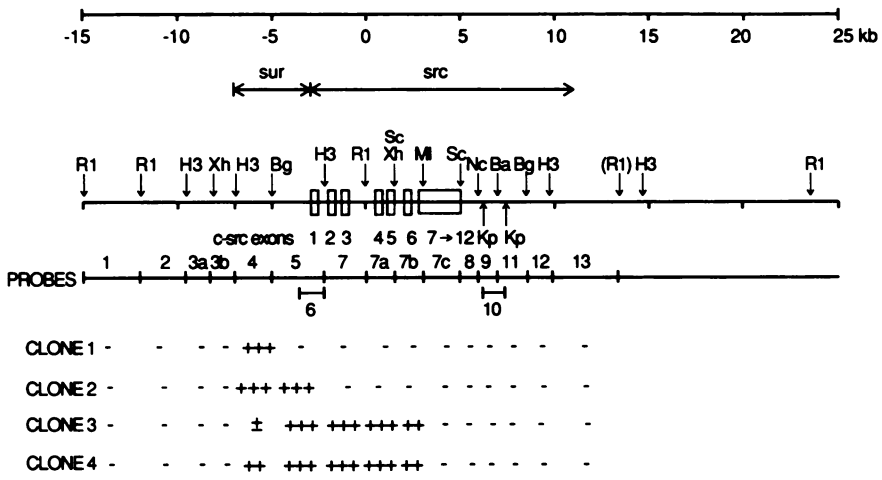

FIG. 2. Structure of the chicken c-*src* locus with the restriction map used to generate various genomic probes. This figure is a summary of the sequence content of the various cDNA clones from several Southern blot analyses. Exons 7 through 12 (7 → 12) are boxed together without introns for convenience. The restriction sites are taken from published studies (56, 65). The *Eco*RI site shown in parentheses is the artificial cloning site used for the c-*src* DNA. The scale (in kilobase pairs) is shown on top. The boundaries for the *src* region and the *src* upstream region denoted as sur, are marked. The notation used to denote various genomic probes is adapted from that of Wang et al. (69). Results of the hybridization studies are indicated below each probe. Abbreviations used for the restriction enzymes are as follows: R1, *Eco*RI; H3, *Hin*dIII; Bg, *Bgl*II; Ml, *Mlu*I; Xh, *Xho*I; Sc, *Sac*I; Nc, *Nco*I; Ba, *Bam*HI; Kp, *Kpn*I.

At this point, it is not clear whether the 5′ sequence of the 3-kb mRNA is colinear with the sequences 5′ to exon 1 in the c-*src* gene. Hence, it became important to determine whether the sequences of the 3-kb mRNA were interdigitated in the c-*src* gene or whether they represented the 5′ neighboring gene of c-*src* and the 3′ c-*src* sequences in the 3-kb mRNAs are merely the result of an aberrant readthrough problem of gene compaction. Thus, mapping of the position of the 5′ end of the c-*src* 4-kb mRNA relative to that of the 3-kb mRNA was performed.

To obtain the most 5′ cDNA clones of the 4-kb c-*src* mRNA, a recently described procedure for the rapid amplification of cDNA ends was used (17). A typical 5′ cDNA clone obtained by this method and confirmed by sequencing to harbor c-*src* exon 1 and its upstream sequences was used as a probe in the mapping of the 5′ end of the 4-kb CEF mRNA. A detailed study of the cDNA of the 4-kb mRNA and its most 5′ sequences will be described elsewhere. This 5′ cDNA probe was hybridized to a panel of c-*src* genomic DNA fragments including probes 1 through 5 (Fig. 2). The most 5′ cDNA clone of the 3-kb mRNA (clone 1 in Fig. 1) was mapped in a parallel analysis. Results obtained with the 5′ cDNA probe of the 4-kb mRNA are shown in Fig. 3A.

As expected, the cDNA probe hybridized to c-*src* probe 5, which contained exon 1 (lane 7 in Fig. 3A). The only other c-*src* DNA detected was c-*src* probe 2 (lane 2 in Fig. 3A), indicating that the sequences 5′ to the c-*src* exon 1 in this cDNA clone map to a region around 8.0 kb upstream from c-*src* exon 1. Figure 3B shows that the most 5′ cDNA probe of the 3-kb mRNA hybridized strongly with c-*src* probe 4 (lane 5). The bands seen in lane 6 represent a partial *Bam*HI digest of c-*src* probe 4, and those in lane 7 represent the probe 5 DNA contaminated with the probe 4 DNA. Therefore, the most 5′ cDNA sequence of the 3-kb mRNA mapped within the c-*src* probe 4 region, a region 4 to 5 kb upstream from the previously defined c-*src* exon 1.

These results imply that the two types of c-*src* mRNAs may initiate at different sites, with the start site for the 3-kb RNA downstream from that for the 4-kb RNA (Fig. 4). Definite proof for separate initiations will await the identification of the cap sites and characterization of the promoters. However, it is clear that the two mRNA sequences are interdigitated in the c-*src* locus. It follows that they have to be generated by alternative splicing and/or differential initi-

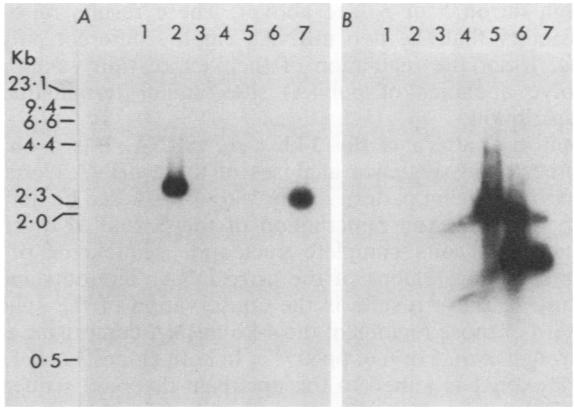

FIG. 3. Mapping the 5′ origins of the 4-kb and 3-kb c-*src* mRNAs. Southern blot hybridization of a 0.5-kb 5′ CEF c-*src* cDNA clone generated by the rapid amplification of cDNA ends method (A) and the muscle cDNA clone 1 (B) to a set of a c-*src* DNA fragments (100 ng of each) run on duplicate gels. The c-*src* DNA fragments (see Fig. 2) selected for this analysis were probe 1 (lane 1), probe 2 (lane 2), probe 3a (lane 3), probe 3b (lane 4), probe 4 (lane 5), a partial *Bam*HI digest of probe 4 (lane 6), and a *Bam*HI digest of probe 5 (lane 7). The molecular size markers are indicated on the left. The preparation of probe 5 contained c-*src* probe 4 as a contaminant. c-*src* probe 5 has no *Bam*HI site, whereas c-*src* probe 4 has a single *Bam*HI site (69). The upper bands in lanes 6 and 7 of panel B represent the uncut probe 4, and the lower bands represent the *Bam*HI fragment derived from probe 4. The difference in the signal ratio between the upper and lower bands in lanes 6 and 7 reflects the different degrees of *Bam*HI digestion. The signal in lane 7 of panel B is solely due to contaminating probe 4 DNA. The 2.4-kb probe 5 DNA was not detected by this 5′ muscle cDNA.

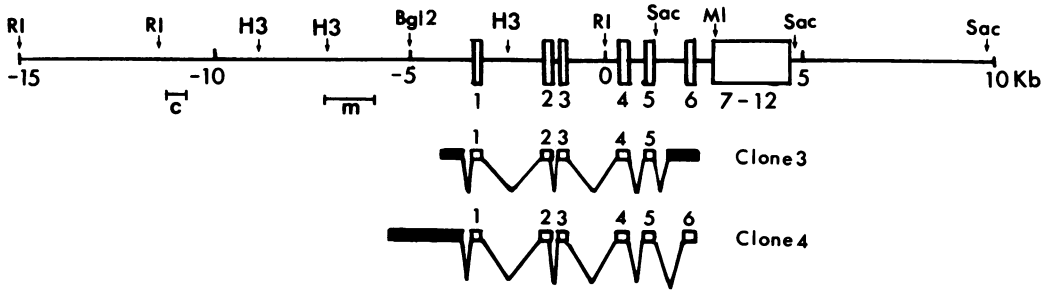

FIG. 4. Schematic patterns of RNA splicing of muscle c-*src* mRNAs. cDNA clones 3 and 4 are depicted. The sequences upstream of exon 1 (in clones 3 and 4) and sequences downstream of exon 5 (in clone 3) are depicted in continuous solid boxes. The punctuation of exons for those sequences is not known. The genomic regions from which those divergent sequences are derived are shown directly above these boxes. However, the genomic and cDNA sequences shown are not strictly in scale. The 3′ diverged sequences for clone 3 are derived from either intron 5 or intron 6, as determined by hybridization analysis. Bars marked c (for CEF *src* mRNA) and m (for muscle *src* mRNA) represent the boundaries of the restriction fragments to which the corresponding 5′-most cDNA clones hybridized. RI, *Eco*RI; H3, *Hin*dIII; Bgl2, *Bgl*II; Sac, *Sac*I; MI, *Mlu*I.

ation of transcription. Our earlier results demonstrated that the 4-kb mRNA in skeletal muscle was detectable only in chicken embryos up to 18 days old, at which time the 3-kb mRNA appeared and persisted thereafter (69). Therefore, this alternative splicing and the presumed differential initiation of transcription must be under tissue-specific and developmental regulation.

With respect to the 3′ ends of the c-*src* mRNAs, our previous data suggested that there exists a 3′ 1.9-kb noncoding sequence for the 4-kb CEF c-*src* mRNA (69). Our cDNA sequence of the 4-kb mRNA confirmed those data and further located the poly(A) signal at 1.9 kb downstream of the last coding exon (exon 12) of the mRNA. Details of this study will be described elsewhere. By contrast, the most 3′ sequence of the cDNA (clone 3) of the 3-kb mRNA mapped within intron 5 or 6 (see above). These results raise the possibility that the two mRNAs employ different poly(A) sites. If so, the regulation of their expression would also involve a choice of poly(A) sites and/or termination of transcription.

Splicing pattern of the 3-kb c-*src* mRNA. Based on the Southern and sequence analyses of these cDNA clones, a tentative splicing pattern of the 3-kb mRNA can be derived (Fig. 4). The exon punctuation of the 5′ and 3′ diverged sequences awaits complete nucleotide sequencing of the corresponding regions of the c-*src* DNA. The outstanding feature of these results is the conservation of the splicing pattern of those regions of the 4-kb mRNA comprising exon 1 through exon 5 or 6 of $pp60^{c\text{-}src}$. In both clones 3 and 4, the c-*src* exon 1 is joined to the upstream diverged sequences precisely at the splice junction identified for that exon (55, 66). On the 3′ side of clone 3, the sequence diverges precisely at the splice junction for exon 5. In clone 4, the 3′ sequence terminates at the end of exon 6. This difference at their 3′ ends most likely accounts for the observed heterogeneity in muscle c-*src* mRNAs (three species of RNA of 2.8, 3.0, and 3.2 kb in size [69]).

The bulk of the 3-kb mRNA represented by the 5′ diverged sequences was derived from c-*src* probe 4 and 5 regions of the c-*src* locus (Fig. 2). These regions are boxed together in Fig. 4, as the exon-intron structures in these regions are not known. For the same reason, the 3′ diverged sequence of clone 3 which is derived from intron 5 or 6 defined by a 1.3-kb *Sac*I-*Mlu*I fragment is also represented by a single box.

Nucleotide sequences of the cDNA clones. A 3,074-nucleotide sequence was composed from the overlapping cDNA clones 1, 2, and 3. When the sequence was translated in all three reading frames, three stretches of long open reading frames became apparent when the frame 1 sequence was considered (Fig. 1, open boxes). Two other reading frames encoded only short stretches of amino acid sequences. Frame 1 also conformed to the amino acid sequence of $pp60^{c\text{-}src}$ exons 1 through 5. However, the three stretches were separated by one or multiple termination codons. Repeated sequencing with independent cDNA clones in both orientations confirmed the presence of those termination codons. Thus, in the 3-kb c-*src* RNA, the $pp60^{c\text{-}src}$ coding exons are located on the 3′ side of the molecule. The entire sequence is shown in Fig. 5.

The first open reading frame used the ATG codon at position 724 of our cDNA sequence. This initiation codon agrees with the rule for efficient initiation of translation for most eucaryotic mRNAs (36). Moreover, it is similar to initiator codons for chicken, rats, and human skeletal muscle proteins, namely, it has a purine at position −3 and a pyrimidine at position +4 (36). This open reading frame codes for 219 amino acid residues, with a calculated molecular mass of 24 kilodaltons (kDa). The second open reading frame started 84 nucleotides downstream from the end of the first frame and coded for 182 amino acids. However, its initiation codon is in a poor context for initiation of protein synthesis (35). The third open reading frame contained the first five exons of $pp60^{c\text{-}src}$ and coded for 193 amino acid residues, ending at a stop codon in the 3′ diverged sequence 35 nucleotides downstream from the end of exon 5.

Based on the dogma of initiation of protein synthesis in eucaryotic mRNAs, the first open reading frame is most likely the genuine coding sequence, whereas the second and third may be silent noncoding sequences. The predicted polypeptide from the first open reading frame contained a potential N-linked glycosylation site at residue 4 and two stretches of hydrophobic sequences, from residues 62 to 72 and from residues 193 to 208 (highlighted by dashed lines in Fig. 5), with a hydrophobicity index of greater than 1.5 (38). At present, it is not clear whether these hydrophobic sequences serve as signal peptide or transmembrane domains. Another characteristic feature is that the predicted polypeptide was somewhat rich in cysteine (17 out of 219 amino acids) and proline (26 out of 219 amino acids) residues.

Search for homologies. In an attempt to identify sequence homologies between the *sur* gene product and other proteins, a computer search of the Dayhoff, Doolittle and translated GenBank sequences was performed. Only the

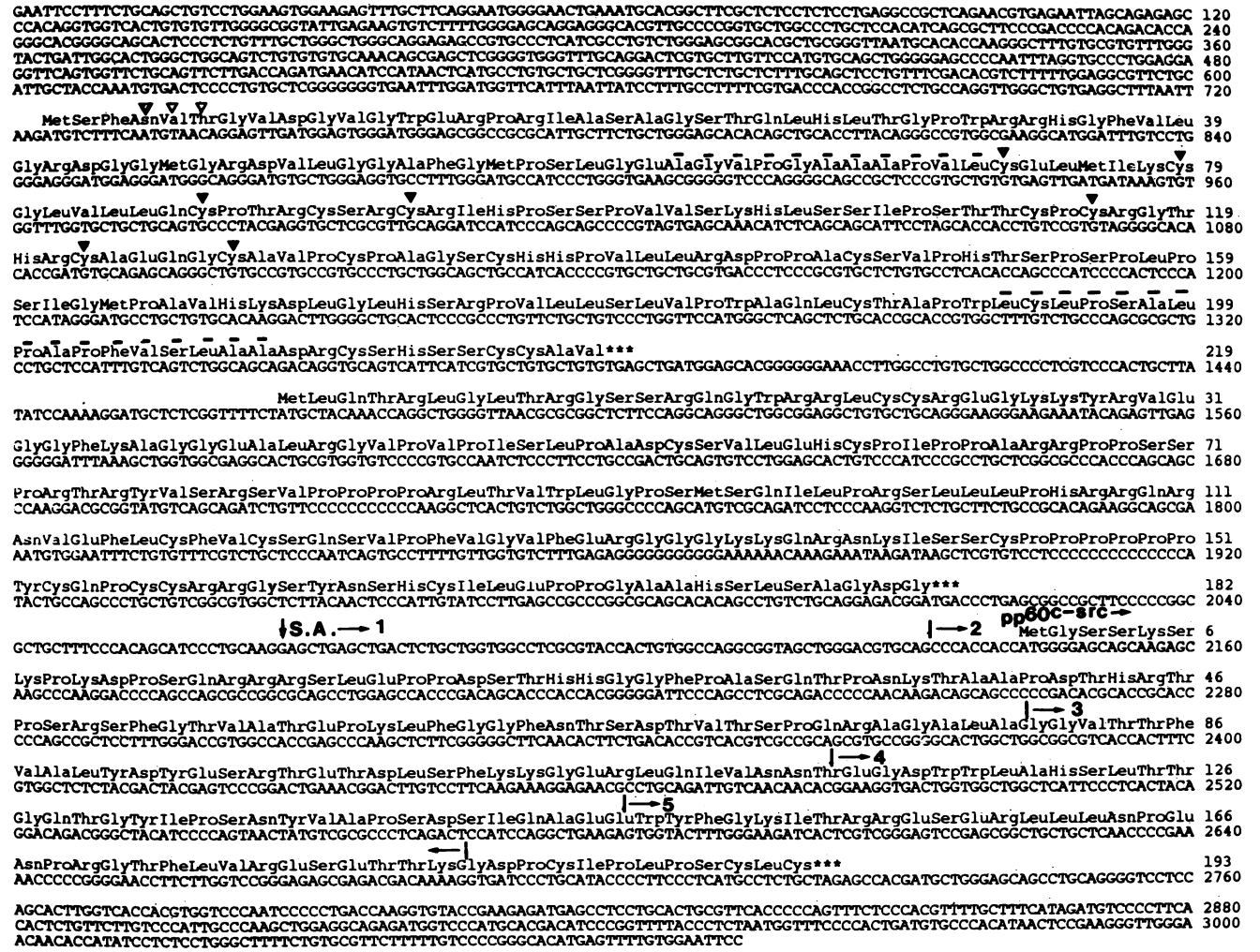

FIG. 5. Nucleotide sequence of the muscle 3-kb c-*src* mRNA. The three open reading frames with their predicted polypeptides are shown. The numbering of amino acids beginning with the initiator methionine for each coding region is shown above the numbering for the nucleotides. The cysteine residues showing homology to mouse EGF precursor are marked by solid triangles. The stop codons are shown by asterisks (***). The amino acids comprising the two hydrophobic regions within the *sur* sequence are overlined. A putative consensus sequence for glycosylation at the beginning of the *sur* sequence is marked by open triangles. The pp60$^{c\text{-}src}$ sequences which form the third open reading frame are shown, with exons 1 through 5 bounded by vertical bars. The position corresponding to the splice acceptor site (S.A.) used for pp60$^{c\text{-}src}$ exon 1 is shown (66).

carboxyl two-thirds of the protein revealed some homology with a class of mammalian proteins that contain sequence stretches similar to epidermal growth factor (EGF), especially the mouse EGF precursor. The spatial distribution of cysteine residues in the region covering amino acids 875 to 931 in the EGF precursor was quite similar to that in the *sur* gene product, between amino acids 73 and 127. The mature EGF molecule, 53 amino acids in size, is encoded within a 1,217-amino-acid precursor in the mouse (24). This large protein, which has all the landmark features of a typical transmembrane protein, displays nine "EGF-like" repeats in its extracellular domain, the ninth of which gives rise to mature EGF by a scheme of extracellular proteolytic processing.

The observed sequence homology with the *sur* gene product covered the fifth, sixth, and seventh EGF-like repeats in the EGF precursor. Although the conservation of the cysteine placement is the salient feature between these two proteins, 10 additional amino acids were also conserved, accounting for an overall homology of 22% in the regions compared. However, the most characteristic C-X-C-X-X-G-F/Y-X-G-X-X-C sequence, where X indicates any amino acid residue, common to EGF and most other EGF-like repeats was not present in the cysteine-rich region of the *sur* gene product. Thus, the *sur* protein is probably not a member of the EGF family. The significance of the similarity in the spatial arrangement of cysteine residues is at present unclear.

In vitro translation and characterization of the *sur* coding sequence. To analyze the nature of the polypeptide encoded by the 3.0-kb mRNA, a nearly fully-length cDNA was constructed by joining the sequences of clones 2 and 3 at their unique *Bgl*II site (Fig. 1). The structure of the construct is depicted in Fig. 6A. The polypeptide produced from the RNA transcript obtained from this construct was analyzed by SDS-PAGE and is shown in Fig. 7A. Even though the RNA transcript was 2.6 kb in size, only a small protein of about 24 kDa was produced (Fig. 7A, lane 2). The in

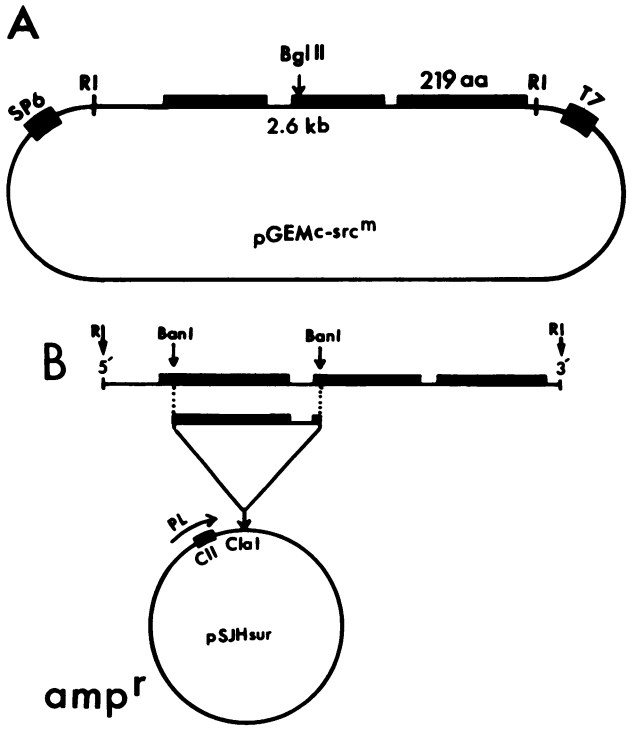

FIG. 6. Transcription and expression vectors of the muscle c-*src* cDNA. (A) The pGEM c-*src*^{m} construct contains 2.6 kb of the muscle c-*src* cDNA sequences, including all three open reading frames. The *sur* coding sequence of 219 amino acids (aa) is on the right, in the correct orientation for T7 RNA polymerase. The cloning sites used were *Eco*RI (RI) and *Bgl*II. Other details are described in Materials and Methods. (B) Structure of the construct (pSJH *sur*) used to express the *sur* coding sequence in bacteria. λ *p*L, the *c*II promoter, is shown with a thin curved arrow. The details of the construction are described in Materials and Methods.

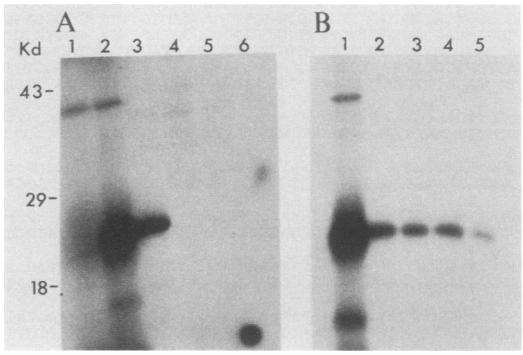

FIG. 7. In vitro expression of the *sur* coding sequence. (A) mRNAs transcribed from pGEM c-*src*^{m} were translated in vitro with rabbit reticulocyte lysates. [^{35}S]methionine-labeled translation products were analyzed by SDS-PAGE either before (lane 2) or after (lane 3) immunoprecipitation with the anti-*sur* antiserum. In parallel, immunoprecipitation was performed with monoclonal antibodies 327 (lane 4) and GD11 (lane 5) and with preimmune serum (lane 6). Molecular mass standards are indicated (in kilodaltons). A faint band of about 40 kDa in lanes 1 and 2 is an endogenous in vitro translation product. Lane 1 represents the translated protein obtained in the absence of exogenously added RNAs. (B) The [^{35}S] methionine-labeled translation products were analyzed by SDS-PAGE either before (lane 1) or after (lane 2) immunoprecipitation with anti-*sur* sereum. In parallel, equivalent amounts of immunoprecipitates were treated with alkaline phosphatase (lane 3), potato acid phosphatase (lane 4), and RNases A, T_1, and T_2 (lane 5). The amount of immunoprecipitate used for RNase treatment represented one-third of the amount used for phosphatase treatment. See Materials and Methods for details.

vitro-translated product often appeared as a broad band. The reason for this diffuse pattern is not yet clear. One possibility is that there may be some premature terminations in the translation. It is also possible that the second open reading frame in the transcript, coding for a polypeptide of 182 amino acids, is translated to a certain extent.

To characterize the in vitro-translated protein product further, an immunoprecipitation analysis was carried out with a polyclonal antiserum raised against a bacterially produced *sur* polypeptide (see below). Upon treatment with anti-*sur* antiserum, the upper portion of the in vitro-translated product was immunoprecipitated, resulting in a sharp band (Fig. 7A, lane 3), but it was not precipitated by the preimmune serum (Fig. 7A, lane 6). The in vitro product was of the expected size of the protein encoded by the first open reading frame of 219 amino acids.

To investigate whether the third open reading frame of the cDNA coded for a *src*-related polypeptide, two anti-*src* antibodies were used in the immunoprecipitation. No proteins could be precipitated by monoclonal antibody 327 (Fig. 7A, lane 4), which recognizes the amino-terminal sequences of pp60$^{c\text{-}src}$ (44), or GD11 (lane 5), which recognizes the epitope(s) within amino acids 82 to 169 of pp60$^{c\text{-}src}$ and pp60$^{v\text{-}src}$ (51). These results suggest that the open reading frame corresponding to pp60$^{c\text{-}src}$ exons 1 to 5 is not translated in our in vitro experiments. However, at this point, we cannot rule out the possibility that the 5′ portion of pp60$^{c\text{-}src}$ is expressed in vivo.

To investigate the possibility that the *sur* protein could be phosphorylated, we treated the immunoprecipitate with calf intestinal alkaline phosphatase (Fig. 7B, lane 3) and with potato acid phosphatase (Fig. 7B, lane 4) under appropriate conditions. The results showed that there was no further sharpening or alteration in mobility of the *sur* protein due to dephosphorylation, suggesting that *sur* protein was not extensively phosphorylated.

To check the possibility that some ^{35}S-labeled tRNAs could be nonspecifically precipitated along with *sur* protein, we digested a portion of the immunoprecipitate with a mixture of RNases A, T_1, and T_2. The results (Fig. 7B, lane 5) indicated that there was no detectable contamination of the immunoprecipitated *sur* band by ^{35}S-labeled tRNA and there was no further sharpening of the *sur* band. The lighter band reflected the fact that only one-third of the product was analyzed for RNase digestion.

As described above, the predicted 24-kDa *sur* protein contained two stretches of hydrophobic amino acids. To investigate the possibility that these sequences play a role in the processing and localization of the *sur* protein, we translated the full-length cDNA transcript in the presence of canine pancreas microsomal vesicles. The results showed that there was no detectable processing of the translated product. To investigate the possible location of the *sur* gene product, sedimentation analyses were performed after exposing the product of 0.5 M potassium acetate, which is capable of removing peripheral membrane proteins without breaking open the membrane vesicles (59). More than 95% of the processed product was found to be associated with membranes even after potassium acetate extraction (data not shown).

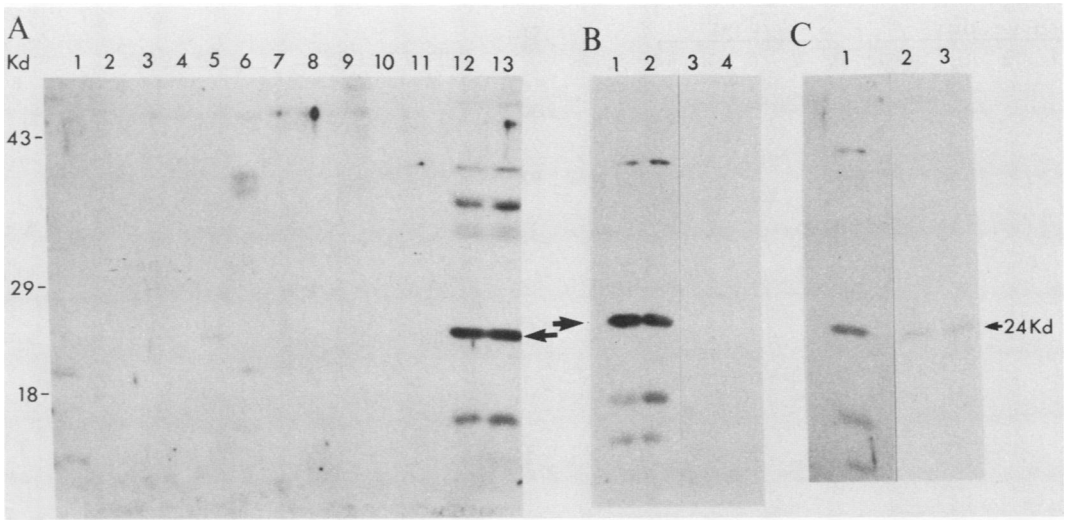

FIG. 8. Expression of *sur* protein in muscle tissue. (A) 50 μg of proteins from each tissue (from 4-day-old chickens) was analyzed by 12% SDS–PAGE followed by immunoblotting and treatment with anti-*sur* serum. The various tissues analyzed were CEF (lane 1); gonad (lane 2); spleen (lane 3); liver (lane 4); pancreas (lane 5); brain (lane 6); bursa (lane 7); lung (lane 8); intestine (lane 9); kidney (lane 10); heart (lane 11); breast muscle (lane 12); and leg muscle (lane 13). (B) An independent analysis showing the muscle-specific proteins in breast muscle (lane 1) and leg muscle (lane 2). Lanes 3 and 4 represent the corresponding pattern obtained with the preimmune serum. (A and B) Molecular mass markers (in kilodaltons) are shown on the left. Arrows point to the 24-kDa *sur* protein. (C) Pattern obtained when an extract prepared from breast muscle tissue (lane 1) was run on the same gel along with two independent immunoprecipitations (lanes 2 and 3) of the ^{3}S-labeled in vitro translation product. The portion of the gel containing lane 1 was processed by Western blotting, and the gel portion containing lanes 2 and 3 was dried directly. Both portions were reassembled later and exposed to X-ray film. Arrows show the 24-kDa band comigrating in lanes 1, 2, and 3.

The orientation of the *sur* protein with respect to the membranes was examined by protease protection experiments after synthesis in the presence of membranes. A mixture of typsin and chymotrypsin completely digested the *sur* protein in both the presence and absence of 1% Triton X-100, suggesting that no major portion of the *sur* protein traversed the membrane (data not shown). Nevertheless, it is possible that a small portion of the *sur* polypeptide is embedded in the membranes, possibly through hydrophobic interactions. As these in vitro studies are only suggestive, more experiments are needed to assess the exact topography of the *sur* protein with respect to the membranes of muscle cells.

Tissue-specific expression of the *sur* gene product. A thermoinducible expression vector was used to express the first open reading frame coding for the *sur* protein in *E. coli* cells (Fig. 6B). A rabbit polyclonal antiserum raised against the bacterially produced *sur* polypeptide was shown to react with the in vitro-translated product of the *sur* RNA (see above). In order to examine the expression of the *sur* gene product in vivo, various tissue extracts from 4-day-old chickens were prepared and subjected to Western analysis. The results are shown in Fig. 8A.

Of the 13 tissues examined, only tissues of skeletal muscle types (breast muscle and leg muscle) (Fig. 8A, lanes 12 and 13) displayed specific proteins recognized by the antiserum. A major 24-kDa and a few other minor bands were observed specifically in muscle tissue. The 24-kDa protein is in agreement with the expected size of the first open reading frame of the 3.0-kb mRNA and the in vitro-translated product. The other minor species of proteins detected from muscle tissues varied among preparations of muscle extracts and immunoblotting experiments.

Another pattern of the muscle-specific proteins obtained in an independent experiment is shown in Fig. 8B. These proteins were recognized by the anti-*sur* serum (lanes 1 and 2) but not by the preimmune serum (lanes 3 and 4). When the muscle extract and the immunoprecipitated, ^{35}S-labeled *sur* protein from an in vitro translation reaction were run on the same gel and processed in parallel, the muscle-specific in vivo 24-kDa protein (Fig. 8C, lane 1) comigrated with the immunoprecipitated in vitro product (Fig. 8C, lanes 2 and 3). This provides suggestive evidence that the 24-kDa proteins observed in vivo and in vitro are the same. Competition experiments in which increasing amounts of bacterially expressed antigen were used to block the antibody produced a reduction in the signals of all the muscle-specific proteins in Western blot analysis (data not shown).

In an effort of clarify whether the 39-kDa and the 16-kDa proteins which were frequently detected together with the 24-kDa protein were a set of fortuitously cross-reacting proteins or in some manner related to the 24-kDa protein, we immunoaffinity purified the polyclonal anti-*sur* serum by passing it through Sepharose or nitrocellulose cross-linked with *sur* antigen. A significant reduction in the titer was observed after purification, and when this purified antiserum was used in Western blots, the 39- and 16-kDa bands could still be seen along with the 24-kDa protein, but the intensity of the signal on the 24-kDa protein was significantly reduced (data not shown). We conclude that there are multiple species of *sur*-specific antibodies with differing affinities for the antigen. It is possible that the 39- and 16-kDa proteins interact with the low-affinity class and the 24-kDa protein interacts with the high-affinity class of *sur*-specific antibodies which could not be recovered from the immunoaffinity column or were more sensitive to denaturation during the elution process. Therefore, the origin of the 39- and 16-kDa proteins remains an open question.

Expression of *sur* and *src* in chicken muscle during development. Finally, we investigated the pattern of expression of

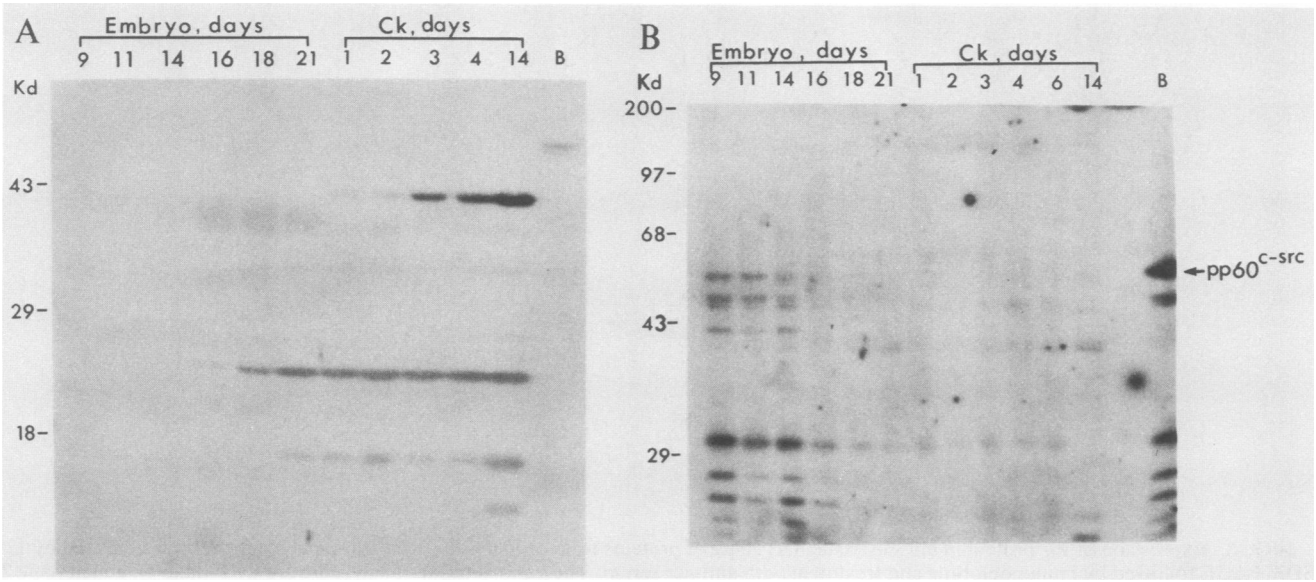

FIG. 9. Kinetics of expression of $p24^{sur}$ and $pp60^{c\text{-}src}$. (A) 50 μg of protein from breast muscle tissue at various stages of embryonic and adult (Ck) life was processed by Western blotting with polyclonal anti-*sur* serum along with a negative control (lane B) of 50 μg of protein from 16-day-old embryonic brain tissue. The details of the blotting procedure are described in Materials and Methods. Molecular mass markers (in kilodaltons) are shown on the left. (B) Pattern obtained when the same muscle extracts used in panel A were processed by Western blotting with monoclonal antibody 327. The molecular mass markers (in kilodaltons) are shown on the left, and the $pp60^{c\text{-}src}$ band in the positive control (16-day-old embryonic brain, lane B) is shown with an arrow. A 52-kDa band below the $pp60^{c\text{-}src}$ band is most likely the frequently observed degradation product of $pp60^{c\text{-}src}$.

$pp60^{c\text{-}src}$ and *sur* in skeletal muscle during various stages of development. Protein extracts were prepared from muscle and brain tissues at various stages of embryonic and adult life of chickens and subjected to Western blot analysis with anti-*sur* or anti-*src* (no. 327) antibody. The results of the kinetic analysis are shown in Fig. 9. The muscle-specific 24-kDa protein was not expressed during early stages of embryonic development (Fig. 9A, 9th, 11th and 14th days). It became visible in the 16-day-old embryonic muscle tissue, and continued to increase up to the 21st day. The time course of appearance of the 24-kDa band correlated very well with that of the 3.0-kb muscle-specific c-*src* mRNA in Northern blot analyses in our previous study (69). It should be noted that the two cross-reacting proteins of 39 and 16 kDa were completely absent when the 24-kDa band appeared in the 16- to 18-day-old embryonic muscle tissue, and the kinetics of appearance of the 39-kDa protein are clearly different from that of the 24-kDa polypeptide. These results suggest that the 39-kDa and the 16-kDa polypeptides are distinctive proteins but are related to the *sur* antigen.

When the same muscle extracts were probed with monoclonal antibody 327, diminishing $pp60^{c\text{-}src}$ expression was seen during early embryonic muscle development (Fig. 9B). The 60-kDa *src* protein (along with its degraded 52-kDa product) was clearly visible in 9-, 11-, and 14-day-old embryonic muscle tissue, and it became undetectable on the 16th day and thereafter. A high level of $pp60^{c\text{-}src}$ was seen when an extract from 14-day-old embryonic brain was run in parallel as a positive control. However, both muscle and brain extracts showed additional cross-reacting bands of low molecular weights, the reason for which is not clear. The disappearance of $pp60^{c\text{-}src}$ with similar kinetics was also documented in our previous study with in vitro kinase assays (69). These results also argue against the possibility that the 5′ portion of $pp60^{c\text{-}src}$ could be expressed from the third open reading frame of the 3-kb RNA in vivo. No muscle-specific $pp60^{c\text{-}src}$-related proteins of any size were detectable after the 16th day of embryonic life. Thus, these studies on the developmental profile of *src* and *sur* proteins demonstrate an interesting inverse relationship between them.

DISCUSSION

We report here the finding of a muscle-specific c-*src* locus-derived mRNA which encodes a small protein of 24 kDa. Generation of this muscle-specific c-*src* mRNA involves alternative RNA splicing, which results in the removal of the TPK domains and incorporation of substantial amounts of sequences upstream from the $pp60^{c\text{-}src}$ coding exons. Expression of the 4-kb versus the 3-kb mRNA in chicken skeletal muscle is under strict developmental control.

Although alternative RNA splicing has been extensively observed for eucaryotic messages, differentially spliced mRNAs leading to entirely different protein products are less common. The synthesis of alternative RNAs from a given gene may involve multiple initiations and multiple poly(A) sites as well as differential splicing (40). Our data imply that generation of the 4-kb and 3-kb mRNAs from the c-*src* gene may involve all three levels of control.

Our mapping of the genomic DNA origins of the most 5′ cDNA sequences for the 4- and 3-kb mRNAs suggests that transcription of these mRNAs may initiate at different promoter sites, with that for the 4-kb RNA upstream from that for the 3-kb mRNA. This result suggests that the *sur* gene transcription unit lies within that for the *src* gene. If so, it would be interesting to see whether the choice of initiation sites for transcription has any role in determining the pattern of splicing and choice of poly(A) sites, which also appear to be different for the two types of c-*src* mRNAs. It is worth noting that the 3- and 4-kb RNAs diverge at their 3′ sides right before the TPK domain (exons 7 to 12 of $pp60^{c\text{-}src}$). It

is tempting to speculate that this punctuation of splicing may reflect separate origins for the TPK and its upstream domains in the formation of the c-*src* gene.

Despite conservation of exons 1 through 5 or 6 of pp60$^{c\text{-}src}$ in the 3-kb mRNA, this region is apparently noncoding in this alternative mRNA. However, this region could still possibly play some role in the processing, stability, or translation of the 3-kb mRNA. These possibilities could be tested by removing those sequences from the cDNA of the 3-kb RNA and testing its efficiency of expression or stability of the message. Also, it would be interesting to know whether this gene organization is conserved in other species. At the protein level, although the anti-*sur* serum detected certain rat- and mouse-specific proteins (data not shown), we do not know whether similar alternatively spliced c-*src* RNA and corresponding *sur* proteins are present in species other than chicken, as observed in this study.

Although the "homology" search indicated that the *sur* protein could be homologous to some cysteine-rich sequences in the mouse EGF precursor protein, it should be noted that there are several cysteines which are not aligned in the two sequences. Particularly, the most conserved EGF motif is absent in the *sur* protein, which weakens the significance of the purported homology.

Our in vitro translation and processing experiments with the *sur* protein suggest that it is an unprocessed membrane-associated protein. Although the protease digestion experiment indicated that no major segment of the polypeptide was sequestered in the membrane vesicles, it remains possible that short stretches of the polypeptide are embedded in the membrane.

Expression of the c-*src* proto-oncogene has been examined extensively (20, 27, 46). Although pp60$^{c\text{-}src}$ is expressed ubiquitously in various chicken tissues at a very low level, its expression in brain is significantly higher, followed by spleen and thymus, whereas its expression in adult chicken muscle is undetectable (7, 23, 54). The 3-kb mRNA was not detected in those earlier studies because in most cases either the TPK domain probe was used for detection or the early embryonic muscle tissue was analyzed.

Expression of the two forms of c-*src* mRNAs is mutually exclusive, and the switch of expression from the 4-kb to the 3-kb RNA is permanent and irreversible, although there appears to be some overlapping during the period of transition, i.e., from the late embryonic stage to hatching (69). Muscle cells are known to generate diversified isomeric contractile proteins through alternative RNA splicing, including myosin heavy chain (52), myosin light chain (49), α-tropomyosin (33), and troponin T (5). However, the *sur* gene product is apparently not a common component of the contractile system, as smooth and cardiac muscle cells do not synthesize this protein. Our earlier work also indicated that the process of myotube formation per se cannot account for the switch of the c-*src* mRNAs. Its appearance during late embryonic stages as well as its membrane association leads us to speculate that *sur* protein may play some role in the formation of neuromuscular junctions and mature postsynaptic apparatus. In this regard, it would be interesting to see how the *sur* gene product localizes with respect to the acetylcholine receptor clusters in muscle cells.

Anthony and co-workers (2) showed that upon infection of a temperature-sensitive mutant of Rous sarcoma virus *ts*NY68, chicken myoblasts could fuse to form multinucleated myotubes at the temperature nonpermissive for transformation. However, those myotubes did not cluster acetylcholine receptor even at the nonpermissive temperature for *ts*NY68 in spite of the presence of active clustering agents. When similar experiments were performed with a transformation-defective mutant of Rous sarcoma virus, the cells behaved like uninfected myotubes, with normal receptor clustering. Their explanation was that the inhibition of clustering might have been mediated by pp60$^{v\text{-}src}$ through its residual kinase activity even at the nonpermissive temperature. Their hypothesis may have some relevance to our observation of the switching-off of the 4-kb c-*src* mRNA if pp60$^{c\text{-}src}$ indeed interferes with the formation of acetylcholine receptor clusters in skeletal muscle cells.

Although our data have not directly related the expression of p24^{sur} to the inhibition of the expression of pp60$^{c\text{-}src}$, the mutually exclusive pattern of the 4-kb and 3-kb c-*src* mRNAs as well as their products strongly suggests that expression of those proteins must be under a rather complicated scheme of control. An intriguing possibility is that p24^{sur} functions as a negative regulator for the expression of pp60$^{c\text{-}src}$, whose presence may be detrimental to mature muscle function. Alternatively, p24^{sur} may play an important role in muscle function and may have nothing to do with the regulation of pp60$^{c\text{-}src}$ expression.

ACKNOWLEDGMENTS

We thank all the members of the Hanafusa laboratory at the Rockefeller University, especially Anindya Dutta, for their help and criticisms. We also thank Song-Muh Jong for several useful discussions and critical reading of the manuscript. We also thank Debkumar Pain and Makato Watanabe of the Blobel laboratory for help with the membrane processing experiments. We acknowledge the gift of monoclonal antibodies 327 from Joan Brugge and EB8 and GD11 from Sarah Parsons and the expression vector pSJH57 from J. H. Chen. Our thanks are also due to Lily Kang during the final stages of this work and to Anita Spingarn and Mirjami Easton for preparation of the manuscript.

This work was supported initially by Public Health Service grants CA-29339 and CA-18213 and later by CA-49400 from the National Institutes of Health.

LITERATURE CITED

1. **Addison, R.** 1987. Secretory protein translocation in a *Neurospora crassa in vitro* system. J. Biol. Chem. **262**:17031–17037.
2. **Anthony, D. T., S. M. Schuetze, and L. L. Rubin.** 1984. Transformation by Rous sarcoma virus prevents acetylcholine receptor clustering on cultured chicken muscle fibers. Proc. Natl. Acad. Sci. USA **81**:2265–2269.
3. **Barnekow, A., and M. Schartl.** 1984. Cellular *src* gene product detected in the freshwater sponge *Spongilla lacustris*. Mol. Cell. Biol. **4**:1179–1181.
4. **Bishop, J. M., and H. E. Varmus.** 1985. Functions and origins of retroviral transforming genes, p. 249–356. *In* R. Weiss, N. Teich, H. E. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Breitbart, R. E., H. T. Nguyen, R. M. Medford, A. Destree, V. Mahdavi, and B. Nadal-Ginard.** 1985. Intricate combinatorial patterns of exon splicing generate multiple regulated troponin T isoforms from a single gene. Cell **41**:67–82.
6. **Brugge, J. S., P. C. Cotton, A. E. Queral, J. N. Barrett, D. Nonner, and R. W. Keane.** 1985. Neurons express high levels of a structurally modified, activated form of pp60$^{c\text{-}src}$. Nature (London) **316**:554–557.
7. **Brugge, J. S., P. Cotton, A. Lustig, W. Yonemoto, L. Lipsich, P. Coussens, J. N. Barrett, D. Nonner, and R. W. Keane.** 1987. Characterization of the altered form of the c-*src* gene product in neuronal cells. Genes Dev. **1**:287–296.
8. **Cartwright, C. A., R. Simantov, P. L. Kaplan, T. Hunter, and W. Eckhart.** 1987. Alterations in pp60$^{c\text{-}src}$ accompany differentiation of neurons from rat embryo striatum. Mol. Cell. Biol. **7**:1830–1840.
9. **Cartwright, C. A., R. Simantov, W. M. Cowan, T. Hunter, and**

W. Eckhart. 1988. pp60$^{c\text{-}src}$ expression in the developing rat brain. Proc. Natl. Acad. Sci. USA **85**:3348–3352.
10. **Chen, J. H.** 1985. The proto-oncogene c-*ets* is preferentially expressed in lymphoid cells. Mol. Cell. Biol. **5**:2993–3000.
11. **Collett, M. S., A. F. Purchio, and R. L. Erikson.** 1980. Avian sarcoma virus-transforming protein pp60$^{v\text{-}src}$ shows protein kinase activity specific for tyrosine. Nature (London) **285**:167–169.
12. **Cooper, J. A., K. L. Gould, C. A. Cartwright, and T. Hunter.** 1986. Tyr-527 is phosphorylated in pp60$^{c\text{-}src}$: implications for regulation. Science **231**:1431–1434.
13. **Cooper, J. A., and C. S. King.** 1986. Dephosphorylation or antibody binding to the carboxy terminus stimulates pp60$^{c\text{-}src}$. Mol. Cell. Biol. **6**:4467–4477.
14. **Courtneidge, S. A., A. D. Levinson, and J. M. Bishop.** 1980. The protein encoded by the transforming gene of avian sarcoma virus (pp60^{src}) and a homologous protein in normal cells (pp60$^{proto\text{-}src}$) are associated with the plasma membrane. Proc. Natl. Acad. Sci. USA **77**:3783–3787.
15. **Coussens, P. M., J. A. Cooper, T. Hunter, and D. Shalloway.** 1985. Restriction of the in vitro and in vivo tyrosine protein kinase activities of pp60$^{c\text{-}src}$ relative to pp60$^{v\text{-}src}$. Mol. Cell. Biol. **5**:2753–2763.
16. **Dale, R. M. K., B. A. McClure, and J. P. Houchins.** 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid **13**:31–40.
17. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85**:8998–9002.
18. **Fults, D. W., A. C. Towle, J. M. Lauder, and P. F. Maness.** 1985. pp60$^{c\text{-}src}$ in the developing cerebellum. Mol. Cell. Biol. **5**:27–32.
19. **Gee, C. E., J. Griffin, L. Sastre, J. Miller, T. A. Springer, H. Piwnica-Worms, and T. M. Roberts.** 1986. Differentiation of myeloid cells is accompanied by increased levels of pp60$^{c\text{-}src}$ protein and kinase activity. Proc. Natl. Acad. Sci. USA **83**:5131–5135.
20. **Gessler, M., and A. Barnekow.** 1984. Differential expression of the cellular oncogenes c-*src* and c-*yes* in embryonal and adult chicken tissues. Biosci. Rep. **4**:757–770.
21. **Golden, A., S. P. Nemeth, and J. S. Brugge.** 1986. Blood platelets express high levels of the pp60$^{c\text{-}src}$-specific tyrosine kinase activity. Proc. Natl. Acad. Sci. USA **82**:852–856.
22. **Golden, A., and J. S. Brugge.** 1988. The *src* oncogene, p. 169–173. *In* E. P. Reddy, T. Curran, and A. Skalka (ed.), The oncogene handbook. Elsevier Science Publishers, New York.
23. **Gonda, T. J., D. K. Sheiness, and J. M. Bishop.** 1982. Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. Mol. Cell. Biol. **2**:617–624.
24. **Gray, A., T. J. Dull, and A. Ullrich.** 1983. Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular-weight protein precursor. Nature (London) **303**:722–725.
25. **Hanafusa, H.** 1986. Activation of the c-*src* gene, p. 100–105. *In* P. Kahn and T. Graf (ed.), Oncogenes and growth control. Springer-Verlag, Berlin.
26. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241**:42–52.
27. **Hoffman, F. M., L. D. Fresco, H. Hoffman-Folk, and B.-Z. Shilo.** 1983. Nucleotide sequences of the *Drosophila src* and *abl* homologs: conservation and variability in the *src* family oncogenes. Cell **35**:393–401.
28. **Hunter, T., and B. M. Sefton.** 1980. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc. Natl. Acad. Sci. USA **77**:1311–1315.
29. **Iba, H., T. Takeya, F. R. Cross, T. Hanafusa, and H. Hanafusa.** 1984. Rous sarcoma virus variants that carry the cellular *src* gene cannot transform chicken embryo fibroblasts. Proc. Natl. Acad. Sci. USA **81**:4424–4428.
30. **Iba, H., F. R. Cross, E. A. Garber, and H. Hanafusa.** 1985. Low level of cellular protein phosphorylation by overproduced pp60$^{c\text{-}src}$. Mol. Cell. Biol. **5**:1058–1066.
31. **Jacobs, C., and H. Ruebsamen.** 1983. Expression of pp60$^{c\text{-}src}$ protein kinase in adult and fetal human tissue. High activities in some sarcomas and mammalian carcinomas. Cancer Res. **43**:1696–1702.
32. **Kamps, M. P., and B. M. Sefton.** 1986. Neither arginine nor histidine can carry out the function of Lys-295 in the ATP-binding site of pp60^{src}. Mol. Cell. Biol. **6**:751–757.
33. **Karlik, C. C., J. W. Mahaffey, M. D. Coutu, and E. A. Fyrberg.** 1984. Organization of contractile protein genes within the 88F subdivision of the D. *melanogaster* third chromosome. Cell **37**:469–481.
34. **Kato, J.-Y., T. Takeya, C. Grandori, H. Iba, J. B. Levy, and H. Hanafusa.** 1986. Amino acid substitutions sufficient to convert the nontransforming pp60$^{c\text{-}src}$ protein to a transforming protein. Mol. Cell. Biol. **6**:4155–4160.
35. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
36. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
37. **Krueger, J. G., E. A. Garber, and A. R. Goldberg.** 1983. Subcellular localization of pp60^{src} in RSV-transformed cells. Curr. Top. Microbiol. Immunol. **107**:51–123.
38. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. **157**:105–132.
39. **Lautenberger, J. A., D. Court, and T. S. Papas.** 1983. High-level expression in *E. coli* of the carboxy-terminal sequences of the avian myelocytomastosis virus (mc29) v-*myc* protein. Gene **23**:75–84.
40. **Leff, S. M., M. G. Rosenfeld, and R. M. Evans.** 1986. Complex transcriptional units and diversity in gene expression by alternative RNA processing. Annu. Rev. Biochem. **55**:1091–1117.
41. **Levy, B. T., L. K. Sorge, A. Meymandi, and P. F. Maness.** 1984. pp60$^{c\text{-}src}$ kinase is in embryonic tissues of chick and human. Dev. Biol. **104**:9–17.
42. **Levy, J. B., H. Iba, and H. Hanafusa.** 1986. Activation of the transforming potential of pp60$^{c\text{-}src}$ by a single amino acid change. Proc. Natl. Acad. Sci. USA **83**:4228–4232.
43. **Levy, J. B., T. Dorai, L.-H. Wang, and J. S. Brugge.** 1987. The structurally distinct form of pp60$^{c\text{-}src}$ detected in neuronal cells is encoded by a unique c-*src* mRNA. Mol. Cell. Biol. **7**:4142–4145.
44. **Lipsich, L. A., A. J. Lewis, and J. S. Brugge.** 1983. Isolation of monoclonal antibodies that recognize the transforming proteins of avian sarcoma viruses. J. Virol. **48**:352–360.
45. **Lynch, S. A., J. S. Brugge, and J. M. Levine.** 1986. Induction of altered c-*src* product during neural differentiation of embryonal carcinoma cells. Science **234**:873–876.
46. **Maness, P. F., L. K. Sorge, and D. W. Fults.** 1986. An early developmental phase of pp60$^{c\text{-}src}$ expression in the neutral ectoderm. Dev. Biol. **117**:83–89.
47. **Maness, P. F., M. Aubry, C. G. Shores, L. Frame, and K. H. Pfenninger.** 1988. c-*src* gene product in developing rat brain is enriched in nerve growth cone membranes. Proc. Natl. Acad. Sci. USA **85**:5001–5005.
48. **Martinez, R., B. Mathey-Prevot, A. Bernards, and D. Baltimore.** 1987. Neuronal pp60$^{c\text{-}src}$ contains a six-amino-acid insertion relative to its non-neuronal counterpart. Science **237**:411–415.
49. **Nebeshima, Y., Y. Kurijama-Fujii, M. Muramatsu, and K. Ogata.** 1984. Alternative transcription and two modes of splicing result in two myosin light chains from one gene. Nature (London) **308**:333–338.
50. **Neckameyer, W. S., M. Shibuya, M. T. Hsu, and L.-H. Wang.** 1986. Proto-oncogene c-*ros* codes for a molecule with structural features common to those of growth factor receptors and displays tissue-specific and developmentally regulated expression. Mol. Cell. Biol. **6**:1478–1486.
51. **Parsons, S. J., D. J. McCarley, V. W. Raymond, and J. T. Parsons.** 1986. Localization of conserved and nonconserved epitopes within the Rous sarcoma virus-encoded *src* protein. J.

Virol. **59**:755–758.
52. **Rozek, C. E., and N. Davidson.** 1983. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell **32**:23–34.
53. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
54. **Schartl, M., and A. Barnekow.** 1984. Differential expression of the cellular *src* gene during vertebrate development. Dev. Biol. **105**:415–422.
55. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
56. **Shalloway, D., A. D. Zelenetz, and G. M. Cooper.** 1981. Molecular cloning and characterization of the chicken gene homologous to the transforming gene of Rous sarcoma virus. Cell **24**:531–541.
57. **Shalloway, D., P. M. Coussens, and P. Yaciuk.** 1984. Overexpression of the c-*src* protein does not induce transformation of NIH 3T3 cells. Proc. Natl. Acad. Sci. USA **81**:7071–7075.
58. **Simon, M. A., B. Drees, T. Kornberg, and J. M. Bishop.** 1985. The nucleotide sequence and the tissue-specific expression of *Drosophila* c-*src*. Cell **42**:831–840.
59. **Singer, S. J., and G. L. Nicolson.** 1972. The fluid mosaic model for the structure of cell membranes. Science **175**:720–731.
60. **Smart, J. E., H. Oppermann, A. P. Czernilofsky, A. F. Purchio, R. L. Erikson, and J. M. Bishop.** 1981. Characterization of sites for tyrosine phosphorylation in the transforming proteins of Rous sarcoma virus ($pp60^{src}$) and its normal cellular homologue ($pp60^{c\text{-}src}$). Proc. Natl. Acad. Sci. USA **78**:6013–6017.
61. **Snyder, M. A., J. M. Bishop, J. P. McGrath, and A. D. Levinson.** 1985. A mutation of the ATP-binding site of $pp60^{v\text{-}src}$ abolishes kinase activity, transformation, and tumorigenicity. Mol. Cell. Biol. **5**:1772–1729.
62. **Sorge, L. K., B. T. Levy, and P. F. Maness.** 1984. $pp60^{c\text{-}src}$ is developmentally regulated in the neural retina. Cell **36**:249–257.
63. **Spector, D. H., H. E. Varmus, and J. M. Bishop.** 1978. Characteristics of cellular RNA related to the transforming gene of avian sarcoma viruses. Cell **13**:381–386.
64. **Sugano, S., M. Y. Stoeckle, and H. Hanafusa.** 1987. Transformation by Rous sarcoma virus induces a novel gene with homology to a mitogenic platelet protein. Cell **49**:321–328.
65. **Takeya, T., H. Hanafusa, R. P. Junghans, G. Ju, and A. M. Skalka.** 1981. Comparison between the viral transforming gene (*src*) of recovered avian sarcoma virus and its cellular homolog. Mol. Cell. Biol. **1**:1024–1037.
66. **Takeya, T., and H. Hanafusa.** 1983. Structure and sequence of the cellular gene homologous to the RSV *src* gene and the mechanism for generating the transforming virus. Cell **32**: 881–890.
67. **Wang, J. Y. J.** 1985. Isolation of antibodies for phosphotyrosine by immunization with a v-*abl*-encoded protein. Mol. Cell. Biol. **5**:3640–3643.
68. **Wang, L.-H., and P. Duesberg.** 1974. Properties and location of poly(A) in Rous sarcoma virus RNA. J. Virol. **14**:1515–1529.
69. **Wang, L.-H., S. Iijima, T. Dorai, and B. Lin.** 1987. Regulation of the expression of proto-oncogene c-*src* by alternative RNA splicing in chicken skeletal muscle. Oncogene Res. **1**:43–59.
70. **Wang, S. Y., W. S. Hayward, and H. Hanafusa.** 1977. Genetic variation in the RNA transcripts of endogenous virus genes in uninfected chicken cells. J. Virol. **24**:64–73.
71. **Watson, C. J., and J. F. Jackson.** 1985. An alternative procedure for the synthesis of double-stranded cDNA for cloning in phage and plasmid vectors, p. 79–88. *In* D. M. Glover (ed.), DNA cloning, vol. 1. IRL Press, New York.