Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5842–5846, September 1981
Microbiology

Cellular homologue (c-*src*) of the transforming gene of Rous sarcoma virus: Isolation, mapping, and transcriptional analysis of c-*src* and flanking regions

(*src*/genome organization)

RICHARD C. PARKER, HAROLD E. VARMUS, AND J. MICHAEL BISHOP

Department of Microbiology and Immunology, University of California, San Francisco, California 94143

Contributed by J. Michael Bishop, May 11, 1981

ABSTRACT **The tumorigenic properties of Rous sarcoma virus are attributable to a 60,000-dalton protein, pp60$^{v\text{-}src}$, encoded by a single viral gene, v-*src*. A homologous gene, c-*src*, that contains the information for a 60,000-dalton protein, pp60$^{c\text{-}src}$, has been identified in all tested vertebrate cells.**

By screening a recombinant DNA library representative of the chicken genome, we isolated two overlapping DNA fragments that contain more than 30 kilobases (kb) of DNA spanning the coding sequences for pp60$^{c\text{-}src}$. This 30-kb region is devoid of moderately or highly repeated sequences and shares homology with the entire viral gene and noncoding sequences 5′ of v-*src*. Although v-*src* has an uninterrupted coding sequence, c-*src* is interrupted by a minimum of seven intervening sequences.

At least two polyadenylylated RNAs seem to be encoded within the DNA we have isolated. The larger RNA, approximately 3.9 kb, is the presumptive c-*src* mRNA; the smaller transcript of about 2 kb hybridizes to DNA sequences several kb from the 3′ end of the v-*src*/c-*src* homology area.

The nucleic acid sequence of the transforming gene of Rous sarcoma virus (RSV), v-*src*, is very similar to a highly conserved sequence found in all vertebrates studied (1, 2). v-*src* encodes a 60,000-dalton protein, pp60$^{v\text{-}src}$,(3) and the cellular homologue, c-*src*, encodes a similar protein, pp60$^{c\text{-}src}$ (4, 5). Differences between viral and cellular *src* proteins are slight (just as the nucleic acid homology would predict) but can be discerned by peptide mapping (4, 5). The significance of these differences is unknown. Normal cells that express low levels of pp60$^{c\text{-}src}$ become tumorigenic when, after infection by RSV, they contain high levels of pp60$^{v\text{-}src}$. Whether the tumorigenicity is due to the difference in protein structure or in the amount of protein in the cell remains to be determined.

To explore the relationship between v-*src* and c-*src* in greater detail, we have isolated the gene that codes for pp60$^{c\text{-}src}$from a recombinant DNA library representative of the chicken genome. We used a recombinant DNA clone of Schmidt–Ruppin A viral DNA SRA-2 (6) as a probe to detect c-*src* sequences and isolated more than 30 kilobases (kb) of chicken DNA containing the c-*src* locus.

In the chicken genome, c-*src* is within a transcriptionally active region of DNA that seems to encode at least two polyadenylylated RNAs. Although less than 1.7 kb of RNA are needed to encode pp60$^{c\text{-}src}$, the only polyadenylylated RNA detectable by hybridization with c-*src* is approximately 3.9 kb long; a cellular RNA of the same size is known to hybridize to v-*src* and is the presumptive mRNA for pp60$^{c\text{-}src}$. A DNA fragment located thousands of bases from the 3′ end of the v-*src*/c-*src* homology area hybridizes to a polyadenylylated RNA of approximately 2 kb. The exact distance between this transcript and the 3′ end of c-*src* RNA is not known.

Like many eukaryotic genes, c-*src* contains intervening sequences that are not responsible for the primary amino acid structure of pp60$^{c\text{-}src}$. A comparison of v-*src* and c-*src* through electron microscopy, nucleic acid sequence determination, and filter hybridization indicated that c-*src* is disrupted by at least seven intervening sequences. The c-*src* locus contained in these clones shares homology with all of v-*src* and sequences flanking the 5′ end of the viral gene.

MATERIALS AND METHODS

Cloning. A recombinant DNA library representative of the chicken genome (7) was a gift from J. Dodgson and D. Engel and was screened as described (8) with probe II. Subcloning of fragments isolated from this library was done with either pBR322 or p480 (a pBR322 derivative containing sequences from herpes simplex DNA and mouse mammary tumor virus DNA) as vectors.

Part of a genomic clone, λCs2, was subcloned into p480 by deleting sequences from the vector that were located between a *Bgl* II site and a *Kpn* I site and replacing them with a 12-kb fragment from the cellular clone that extends from the *Bgl* II site in the chicken DNA insert to a *Kpn* I site in the left arm of Charon 4A.The resulting subclone, pCs100, is a derivative of pBR322 in which the approximately 350 bases of the plasmid between the *Bam*HI and the *Hin*dIII sites are replaced with 2.8 kb of herpes simplex DNA, 9.5 kb of chicken DNA, 2.5 kb of Charon 4A DNA, and 0.8 kb of mouse mammary tumor virus DNA, arranged 5′ to 3′ of c-*src*, respectively.

Restriction Mapping. Restriction endonucleases were purchased from New England Biolabs and Bethesda Research Laboratories and were used as recommended by the vendors. Agarose gel electrophoresis and restriction endonuclease site mapping were done as described (9, 10).

Filter Hybridization. RNA was isolated from a 13-day-old chicken embryo as described by Varmus *et al.* (11), fractionated by formaldehyde/agarose gel electrophoresis, and transferred to nitrocellulose paper (12). Similarly, restriction endonuclease digestion products of DNA were resolved by agarose gel electrophoresis and transferred to nitrocellulose paper (13). The nitrocellulose filters were hybridized to ^{32}P-labeled probes as described (11).

Electron Microscopy. Samples were prepared for electron microscopy as described by Davis *et al.* (14).

Probes. DNA fragments were radioactively labeled with ^{32}P with either reverse transcriptase (RNA-dependent DNA nucleotidyltransferase) (15) or *Escherichia coli* DNA polymerase

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

Abbreviations: RSV, Rous sarcoma virus; kb, kilobase(s)

I (16). Viral probes I–III were made from subclones of a molecular DNA clone of the Schmidt–Ruppin A strain of RSV (6); viral probe IV was made from a DNA clone of part of the B77 strain Rous sarcoma provirus from a transformed rat cell line (11). These probes, defined in relation to the putative start codon for v-*src*, encompass the following regions (v-*src* begins at nucleotide +1 and ends at nucleotide +1591): I, 3′ end of the envelope gene (*env*), envelope/*src* intergenic region, 5′ end of v-*src* (extends from −875 to +685); II, 3′ portion of v-*src*, (extends from +686 to +1556); III, 36 nucleotides from the 3′ end of v-*src*, 3′ noncoding regions to end of viral RNA, 5′ end of RSV RNA (extends from +1557 through the end of the viral RNA and includes 600 bases from the 5′ end of RSV RNA); and IV, 5′ flanking sequences of v-*src*, (extends from −63 to ≈−155).

Probes I, II, and III are defined by *Pvu* II recognition sites.

RESULTS

Clones Containing c-*src*. Two clones, λCs1 and λCs2, that contain the c-*src* locus and flanking regions were isolated and characterized by restriction endonuclease site mapping (Fig. 1). These clones span more than 30 kb of the chicken genome while sharing an ≈4-kb overlap within c-*src*. Variants were formed during propagation of λCs2 as a result of deletions of approximately 1, 5, and 6 kb from the chicken DNA (data not shown).

Comparing v-*src* and c-*src*. With established restriction endonuclease site maps of λCs1 and λCs2, it was possible to compare the viral and cellular genes. DNA fragments of λCs1 and λCs2 formed by cleavage with restriction endonucleases were resolved by agarose gel electrophoresis (Fig. 2*a*), transferred to nitrocellulose filter paper, and hybridized to ^{32}P-labeled probes (13). Initially, three different subclones of SRA-2, defining three regions of the viral genome, were used as probes. These regions are labeled I, II, and III in Fig. 1.

We first attempted to determine the orientation of c-*src* with respect to v-*src* and, by inference, the transcriptional orientation of c-*src*. This was done initially by hybridizing filters with probe I, which represents 5′ flanking sequences and the 5′ 685 nucleotides of v-*src*. This probe recognized two fragments in the restriction endonuclease digests of λCs1 and three in the restriction endonuclease digests of λCs2 (Fig. 2*b*).

To locate the 3′ end and to define the direction of transcription, filters with the same DNA fragments as in Fig. 2*a* were hybridized to probe II. This probe consists entirely of sequences within the 3′-half of v-*src*; it recognized only one fragment in each clone (Fig. 2*c*). That implies that the two fragments in λCs2 that hybridized to probe I but not to probe II must contain the 5′ end of the v-*src*/c-*src* homology in these clones. Additionally, λCs1 must be located on the 3′ end of c-*src*.

These inferences were confirmed by hybridizing filters with probe III (Fig. 2*d*), which contains the last 36 nucleotides of v-*src* and 3′ flanking sequences. This 3′ probe hybridized to λCs1 but not to λCs2. From these data, it is clear that λCs1 contains the most 3′ sequences.

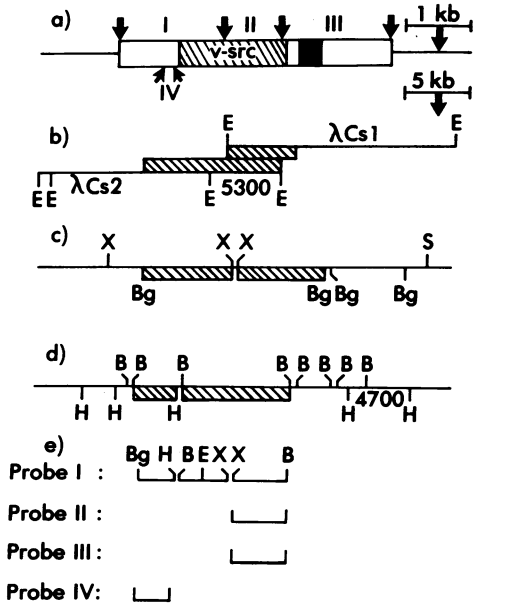

FIG. 1. (*a*) The region of RSV DNA surrounding v-*src*. The shaded box represents the unique sequences from the 3′ and 5′ ends of viral RNA present at both ends of viral DNA. Arrows denote the boundaries of four probes used to characterize cloned cellular DNAs. (*b*) Overlapping regions of cellular DNA inserts in λCs1 and λCs2 were determined by restriction mapping. The *Eco*RI sites (E) found at the ends of each insert were added when the recombinant phage library was made; they are not present in chicken DNA. (*c*) *Sal* I (S), *Xho* I (X), and *Bgl* II (Bg) maps of the chicken DNA in λCs1 and λCs2. The hatched areas indicate restriction endonuclease digest fragments which hybridize with v-*src*. (*d*) *Bam*HI (B), and *Hin*dIII (H) maps as in *c*. (*e*) Each line indicates the chicken DNA fragments that hybridize with the probes in *a*.

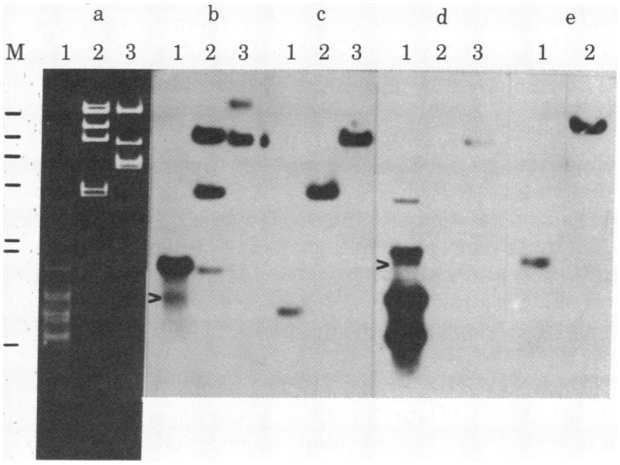

FIG. 2. DNA fragments resolved by agarose gel electrophoresis and detected by short-wave ultraviolet light fluorescence of ethidium bromide-stained DNA (*a*) and by hybridization with ^{32}P-labeled probes: I (*b*), II (*c*), III (*d*), and IV (*e*). Lanes: 1 positive control—SRA-2 digested with *Pvu* II and *Sst* I; 2, λCs2 DNA digested with *Xho* I and *Eco*RI; 3, λCs1 DNA digested with *Xho* I, *Hin*dIII, and *Sal* I. DNA fragments were transferred from agarose gels to nitrocellulose papers that were used for hybridization with probes I–IV. This figure is a compilation of three gels (A, B–D, E). Samples in *e* were resolved on a 1% agarose gel, whereas the samples in *a–d* were resolved on a 0.8% agarose gel. Accordingly, the large band in lane 2 of *e* is not perfectly aligned with that in lane 2 of *b*; they are the same DNA fragment. Lane 3 was not hybridized to probe IV and therefore, is not shown in *e*. In each control (lane 1), probe I should hybridize to the second largest band in this lane, probe II to the fifth largest, and probe III to the fourth and seventh largest. The DNA was not completely digested by *Pvu* II, so large fragments are detected by some probes. All of the probes hybridized as predicted; however, faint cross-hybridization between probe I and part of probe III (the second and fourth largest bands in lane 1) can be seen in *b* and *d*. The fragments that cross-hybridize are noted by horizontal carets (7) in lane 1 of *b* and *d* and contain one copy of an imperfect direct repeat of ≈100 bases that surrounds v-*src* (17). The marker lane (M) shows the location of DNA fragments produced by digestion of λ DNA with *Hin*dIII. Their sizes in kb are: 23.1, 9.7, 6.6, 4.5, 2.5, 2.2, and 0.49 (from top to bottom) (8).

The hybridization of probe III (Fig. 2*d*), which only contains 36 nucleotides of coding sequence, indicated that λCs1 probably includes the 3′ end of $pp60^{c\text{-}src}$ coding sequences. Most or all of the sequences in probe III were absent from λCs2 (Fig. 2*d*, lane 2).

We next attempted to determine whether λCs2 (the clone containing the most 5′ regions of v-*src*/c-*src* homology) included nucleotides encoding the NH_2 terminus of $pp60^{c\text{-}src}$. A probe was prepared that spanned approximately 90 bases beginning 63 nucleotides 5′ of the putative start codon for $pp60^{v\text{-}src}$ and going away from the gene. This probe (number IV) was made from the B77 strain of RSV and hybridized to sequences in λCs2 (Fig. 2*e*). The DNA fragments that hybridized to probe IV were, as expected, also recognized by probe I.

Based upon these hybridization data we can conclude that the entire v-*src* locus, 5′ flanking sequences, and probably 3′ flanking sequences seem to be represented in the 30-kb region defined by λCs1 and λCs2. The restriction endonuclease mapping data also imply that the gene contains intervening sequences (for example, the 0.2 kb *Xho* I fragment does not hybridize to v-*src* probes) that are not present in v-*src* (Fig. 2*c*).

Intron–Exon Spacing in c-*src*. We obtained more detailed data about the structure of c-*src* by analyzing heteroduplexes of DNA molecules by electron microscopy. The 3′ end of λCs2 was subcloned into p480; this subclone, pCs100, extends from the *Bgl* II site that is 5′ of all cellular sequences with homology to v-*src* to a Kpn I site in the left arm of λ Charon 4A. Heteroduplexes of pCs100 with a plasmid containing v-*src* and its 5′ and 3′ flanking regions were analyzed after digesting each plasmid with *Sal* I. The viral subclone contains a 3000-base *Eco*RI fragment inserted into the *Eco*RI site of pBR322.

An electron micrograph of a resulting heteroduplex can be seen in Fig. 3*a*. A tracing of the molecule shows a long stretch of double-stranded DNA at one end and a small amount of duplex structure at the other. These are regions of pBR322 shared by the two molecules. Adjacent to each of these two arms are large single-stranded bubbles predicted from the known structures of the plasmids; they represent regions of the two clones at either the 5′ end (near the short arm of pBR322) or the 3′ end (near the long arm of pBR322) that are not homologous.

The region between these two large bubbles is the area of primary intrest. There are six deletion loops indicating that there are at least six intervening sequences present in this subclone of c-*src* that are not found in v-*src*. However, there must be at least one intervening sequence in addition to those seen in Fig. 3*a*. pCs100 (derived from λCs2) lacks sequences homologous to v-*src* that are only present in λCs1; therefore, v-*src* that is single stranded in this picture can hybridize to the c-*src* locus. Additionally, the length of the large loop that can be seen in Fig. 3 at the 3′ end of pCs100 indicates that it may include some chicken DNA sequences. This suggests that λCs2 ends in an intervening sequence. To confirm this, the nucleotide sequence of the 3′ end of λCs2 has been determined (data not shown), and it is not homologous to v-*src*; thus, the 3′ end of λCs2 is part of a seventh intervening sequence. Therefore, we conclude that c-*src* contains at least seven intervening sequences that are absent from v-*src*.

Stable RNA from c-*src* and Surrounding Regions. The region of the chicken chromosome represented in our clones is comprised of DNA that does not appear to contain any moderately or highly repeated sequences (data not shown). Therefore, DNA fragments from this region seemed suitable as probes in hybridization experiments with cellular RNA.

Cellular polyadenylylated RNA from chicken embryos was separated by formaldehyde/agarose electrophoresis, transferred to nitrocellulose paper, and hybridized to ^{32}P-labeled

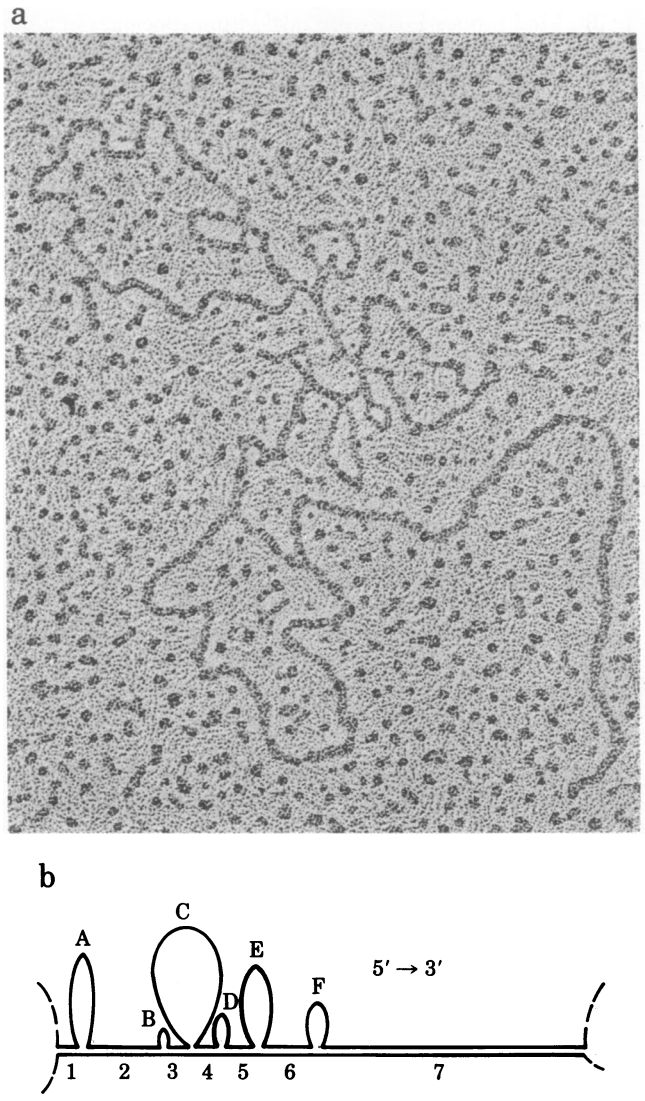

FIG. 3. Electron micrograph of a heteroduplex between plasmids containing v-*src* and most of c-*src* (*a*) and a schematic representation of the introns and exons detected in the micrograph (*b*). The dotted lines at each end of *b* indicate the large bubbles in the micrograph. Six intervening sequences (A–F) and seven regions of homology (1–7) are detected. Their lengths from 5′ to 3′ in kb are: 1, 0.1; 2, 0.27; 3, 0.12; 4, 0.1; 5, 0.12; 6, 0.18; 7, 1.0; A, 1.2; B, 0.2; C, 2.3; D, 0.45; E, 1.1; F, 0.55.

DNA fragments. A single band was detected by hybridization with a viral subclone (probe II) (Fig. 4, lane 1). Its apparent length of 3.9 kb is consistent with published data concerning the size of the putative c-*src* mRNA (18). A band of the same size (and no others) was detected when the 5.3-kb *Eco*RI fragment of λCs2 was used as a probe (Fig. 4, lane 2). This probe contains most of the $pp60^{c\text{-}src}$ coding region and approximately 4 kb of intervening sequences.

A subclone containing the 4.7-kb-long *Hin*dIII fragment of λCs1 (see Fig. 1) hybridized to a smaller transcript of about 2 kb (Fig. 4, lane 3). We assume this transcript is encoded by sequences within this region of the chicken chromosome; however, it is possible that the transcript is encoded elsewhere and that the hybridization detected is due to a small area of homology between the transcript and the probe. This *Hin*dIII fragment is located thousands of bases away from the most 3′ region of v-*src*/c-*src* homology. the existence of this transcript implies that λCs1 contains the 3′ end of the c-*src* locus.

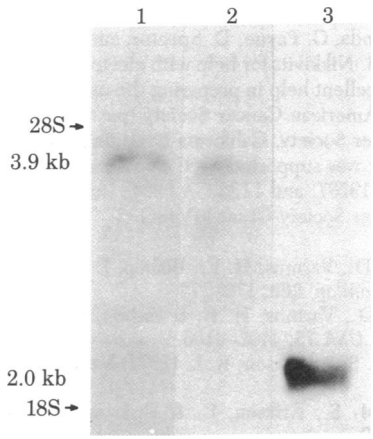

FIG 4. Polyadenylylated chicken embryo RNA was fractionated on a formaldehyde/agarose gel, transferred to nitrocellulose paper, and hybridized to probe II (lane 1), 5.3-kb *Eco*RI fragment of λCs2 containing part of c-*src* (lane 2), and 4.7-kb *Hin*dIII fragment of λCs1 (lane 3). The positions of 18S and 28S ribosomal RNA are indicated by arrows.

DISCUSSION

Are λCs1 and λCs2 Representative of the c-*src* Locus of the Chicken Genome? Although the region of the chicken chromosome included in λCs1 and λCs2 is homologous with the entirety of v-*src*, it is possible that is does not comprise an accurate representation of the c-*src* locus. The restriction endonuclease maps of this region (Fig. 1) can be compared to published hybridization data of genomic blotting of chicken DNA with v-*src* as a probe (19) and to our unpublished results.

The simplest analysis involves the comparison of patterns formed upon digestion with either *Eco*RI or *Bgl* II. The maps in Fig. 1 predict that two chromosomal DNA fragments formed by digestion with *Eco*RI and one fragment formed by digestion with *Bgl* II should hybridize to v-*src*. In fact, an extra fragment is found in each case. This suggests that either the region of the chromosome we have analyzed is missing some of the c-*src* locus, or that there are sequences elsewhere in the genome that are homologous to v-*src*. The latter explanation appears to be correct. If the entire $pp60^{c\text{-}src}$ coding region is represented in λCs1 and λCs2, then sequences elsewhere in the genome that share homology with v-*src* would also share homology with part of the DNA in these clones. We observe such hybridization when λCs2 is used as a probe but not when λCs1 is used as a probe. Therefore, the hybridization to v-*src* is not indicative of an absence of homologous sequences from our clones; instead, it is indicative of another locus that is homologous to part of v-*src*. Modifications of the cellular DNA (i.e., methylation) are unlikely to account for the additional fragments because, with certain enzymes, the additional fragments observed are smaller than those predicted by the maps in Fig. 1.

Which Cellular Locus Encodes $pp60^{c\text{-}src}$? The purpose of this study was to isolate the gene that encodes $pp60^{c\text{-}src}$. The presence of two loci within the chicken chromosome that share homology with v-*src* slightly complicates this problem. We believe that the locus represented by the combination of λCs1 and λCs2 specifies the expression of $pp60^{c\text{-}src}$. We draw this conclusion because this locus is comprised of sequences spanning the entire region of v-*src* (Fig. 2), and because the additional locus hybridizes to v-*src* far less efficiently than does this locus (ref 19; unpublished results).

The majority of the $pp60^{c\text{-}src}$ coding region is represented in the 4-kb segment that is present in both λCs1 and λCs2. If the second locus were responsible for the production of $pp60^{c\text{-}src}$, it would have to include all of these sequences. Because λCs1 only hybridizes to DNA fragments predicted by the map in Fig. 1, a region of at least 17 kb would have to be duplicated at the second locus. In fact, λCs1 DNA hybridizes to a 24-kb *Eco*RI fragment of cellular DNA that would have to be derived from both loci. Such an extensive duplication seems unlikely, so we conclude that $pp60^{c\text{-}src}$ is encoded by the locus presented here.

Is All of c-*src* Within These Clones? Previous data, and the results shown in Fig. 4 suggest that c-*src* mRNA is approximately 3.9 kb; we do not know the length of the primary transcript of the locus. We have not made a cDNA clone of this low-abundance cellular polyadenylylated RNA (18), and we have only defined the boundaries of the locus by comparison with v-*src* and its flanking region. This comparison is limited by the small complexity of the probes (approximately 1900 bases) accounting for about half of the sequences in the putative c-*src* mRNA.

For these reasons, we do not know whether or not the entire c-*src* locus can be found in λCs1 and λCs2. The 3′ end is probably present because we detect an additional, structurally unrelated transcript from sequences 3′ of the v-*src*/c-*src* homology. Based upon hybridization with v-*src* (Fig. 2*e*), these clones contain noncoding sequences flanking the 5′ side of v-*src*; however, we do not know whether or not they contain the site for initiation of transcription of c-*src*.

What Are the Structural Characteristics of c-*src*? During propagation of λCs2, a large number of phage spontaneously delete the 5′ part of c-*src* and flanking noncoding regions. Similar deletions have been seen in the growth of recombinant phage containing retrovirus proviruses (20) and globin genes (21). In those cases, the deletions have been attributed to the presence of direct repeats in the cellular sequences. A similar set of structures may be responsible for the deletions that we observe when propagating c-*src*. The smallest deletion detected, about 1 kb from λCs2, does not delete c-*src* coding sequences.

Based upon the electron microscopy and hybridization data, we can conclude that c-*src* contains at least seven intervening sequences. Because we have only characterized that part of the gene that shares homology with v-*src* and because of the limitations of the techniques employed, it seems possible that there are additonal intervening sequences in this locus.

The most 5′ and 3′ regions of v-*src*/c-*src* homology detected in Fig. 3 are potentially significant. The short stretch of homology at the 5′ -end before the first intron may be within the 5′ untranslated region. It is not possible without nucleic acid sequence determination to define the exact location of initiation of protein synthesis.

At the 3′-end, there is a very large region of v-*src*/c-*src* homology that accounts for over half of the coding sequences of the protein but apparently lacks the carboxy-terminal amino acids (Fig. 2*d* shows that the clone used for electron microscopy fails to hybridize to probe III).

The cellular homologues of the transforming genes of nine retroviruses have now been analyzed for intervening sequences either by filter hybridization to restriction fragments of genomic DNA or by isolation of the specific element. These include the cellular counterparts of the transforming genes of Snyder–Theilen feline sarcoma virus (22): avian erythroblastosis virus, avian myelocytomatosis virus, and avian myeloblastosis virus (unpublished data); Abelson leukemia virus (23); Harvey sarcoma virus (24); simian sarcoma virus (25); and Moloney murine sarcoma virus (26). All except the cellular homologue of the Moloney murine sarcoma virus transforming gene (c-*mos*) contain intervening sequences that are not present in the respective virus.

How Could a Retrovirus Recover c-*src*? Our characterization of c-*src* does not elucidate a mechanism by which a retrovirus could capture this gene. The existence of many intervening sequences suggests that the event is not mediated simply by DNA–DNA recombination. Such a mechanism would be possible, however, if the recombined DNA were transcribed into RNA that was ultimately processed to remove the intervening sequences.

After determining the nucleic acid sequence of v-*src* and the flanking regions from a molecular DNA clone of the Schmidt–Rupin A strain of RSV, Czernilofsky *et al.* (17) noted an imperfect direct repeat (99 out of 126 bases) that flanks v-*src*. This sequence might be involved in capture of the cellular gene by a retrovirus by allowing for homologous recombination between a copy of this sequence found in a retrovirus and a copy near a cellular gene. However, we did not detect the existence of a similar sequence in c-*src*. If the repeat were present in λCs1 or λCs2, it should have been detected in the experiments shown in Fig. 2 *b* and 2*d*. These experiments involved the hybridization of probes from either the 5′ end of v-*src* probe I (Fig. 2*b*) or the 3′ end of v-*src* probe I (Fig. 2d); both of these probes contain one copy of the imperfect direct repeat. The failure of the probe I to hybridize to the 3′ end of c-*src* and the failure of the probe III to hybridize to the 5′ end of c-*src* suggests that the repeat is not present at either the 5′ end or the 3′ end of c-*src*, as it is represented in λCs1 and λCs2. Lane 1 of each of these figures is a positive control for the experiments. It contains viral DNA digested with the same enzyme (*Pvu* II) that was used to isolate the probes. The hybridization seen is a result of the probe hybridizing to the fragments from which the probe was made and also hybridizing to an additonal fragment of viral DNA that contains a copy of the direct repeat. This cross-hybridization implies that we would have detected the presence of the repeat in the cellular DNA.

Will c-*src* Be Helpful in Understanding Oncogenesis? Although tumor viruses can cause many types of cancers, it is clear that these tumors are only a fraction of the many types of neoplasia that affect vertebrates (27). Retroviruses help to define a class of conserved cellular genes that may be crucial to the life cycle of the animal and are oncogenic when expressed in an aberrent fashion.

It is possible that these genes need not be captured by a retrovirus (and we assume that many have not been captured by retroviruses) in order to be tumorigenic. Conceivably, a variety of physical and chemical carcinogens are tumorigenic simply because they allow for aberrant expression of normal cellular genes. For example, recent data suggest that avian leukosis virus causes lymphomas by altering the expression of the cellular homologue of the avian retrovirus MC-29 transforming gene (28).

Having isolated c-*src* it may now be possible to determine whether or not the normal protein, $pp60^{c\text{-}src}$, is capable of transforming avian and mammalian cells. Similar experiments with c-*mos* (26) and the endogenous homologue of Harvey sarcoma virus (c-*ras*) (24) indicate that they are capable of transforming mammalian cells.

We thank B. Baker, J. Majors, and G. Mardon for generous gifts of reagents; T. Gonda, G. Payne, D. Spector, and R. Swanstrom for helpful discussion; B. Nikkivits for help with electron microscopy; and Bertha Cook for excellent help in preparing the manuscript. R.P. was supported by the American Cancer Society (postdoctoral fellow) and the American Cancer Society, California Division (senior postdoctoral fellow). This work was supported by U.S. Public Health Service Grants CA 12705, CA 19287, and 1T32 CA 09043 (training grant) and by the American Cancer Society Grant MV48G.

1. Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. (1976) *Nature (London)* **260**, 170–173.
2. Spector, D., Varmus, H. E. & Bishop, J. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4102–4106.
3. Brugge, J. S. & Erikson, R. L. (1977) *Nature (London)* **269**, 346–347.
4. Collett, M. S., Erikson, E. & Erikson, R. L. (1978) *Cell* **15**, 1363–1370
5. Oppermann, H., Levinson, A. D., Varmus, H. E., Levintow, L. & Bishop, J. M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1804–1808.
6. DeLorbe, W. J., Luciw, P. A., Varmus, H. E. & Bishop, J. M. (1980) *J. Virol.* **36**, 50–61.
7. Dodgson, J. B., Strommer, J. & Engel, J. D. (1979) *Cell* **16**, 879–887.
8. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
9. Parker, R. C. & Watson, R. M. (1977) *Nucleic Acids Res.* **4**, 1291–1299.
10. Parker, R. C. & Seed, B. (1980) *Methods Enzymol.* **65**, 358–363.
11. Varmus, H. E., Quintrell, N. & Ortiz, S. (1981) *Cell*, in press.
12. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
13. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
14. Davis, R. W., Simon, M. & Davidson, N. (1971) *Methods Enzymol.* **21**, 413–428.
15. Shank, P. R., Hughes, S. H., Kung, H. J., Majors, J. E., Quintrell, N., Guntaka, R. V., Bishop, J. M. & Varmus, H. E. (1978) *Cell* **15**, 1383–1395.
16. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184–1188.
17. Czernilofsky, A. P., Levinson, A. D., Varmus, H. E., Bishop, J. M., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **287**, 198–203.
18. Spector, D. H., Baker, B., Varmus, H. E. & Bishop, J. M. (1978) *Cell* **13**, 381–386.
19. Hughes, S. H., Vogt, P. K., Stubblefield, E., Robinson, H., Bishop, J. M. & Varmus, H. E. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **44**, 1077–1089.
20. Vande Woude, G. F., Oskarsson, M., Enquist, L. W., Nomura, S., Sullivan, M. & Fishcinger, P. J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4464–4468.
21. Lauer, J., Shen, C. K. J. & Maniatis, T. (1980) *Cell* **20**, 119–130.
22. Franchini, G., Even, J., Sherr, C. J. & Wong-Staal, F. (1981) *Nature (London)* **290**, 154–157.
23. Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. (1980) *Cell* **22**, 777–785.
24. DeFeo, D., Gonda, M. A., Young, H. A., Chang, E. H., Lowy, D. R., Sednick, E. & Ellis, R. W. (1981) *Proc. Natl. Acad. Sci. USA*, in press.
25. Dalla Favera, R., Gelmann, E. P., Gallo, R. C. & Wong-Staal, F. (1981) *Nature (London)* **292**, 31–35.
26. Oskarsson, M., McClements, W. L., Blair, D. G., Maizel, J. V. & VandeWoude, G. F. (1980) *Science* **207**, 1222–1224.
27. Pimentel, E. (1979) *Biochim. Biophys. Acta* **560**, 169–216.
28. Hayward, W. S., Neel, B. G. & Astrin, S. M. (1981) *Nature (London)* **290**, 475–480.