MOLECULAR AND CELLULAR BIOLOGY, Nov. 1981, p. 1024-1037 Vol. 1, No. 11
0270-7306/81/111024-14$02.00/0

Comparison Between the Viral Transforming Gene (*src*) of Recovered Avian Sarcoma Virus and Its Cellular Homolog

T. TAKEYA* AND H. HANAFUSA
The Rockefeller University, New York, New York 10021
R. P. JUNGHANS, G. JU, AND A. M. SKALKA
The Roche Institute of Molecular Biology, Nutley, New Jersey 07110

Received 22 May 1981/Accepted 21 July 1981

Recovered avian sarcoma viruses are recombinants between transformation-defective mutants of Rous sarcoma virus and the chicken cellular gene homologous to the *src* gene of Rous sarcoma virus. We have constructed and analyzed molecular clones of viral deoxyribonucleic acid from recovered avian sarcoma virus and its transformation-competent progenitor, the Schmidt-Ruppin A strain of Rous sarcoma virus. A 2.0-megadalton *Eco*RI fragment containing the entire *src* gene from each of these clones was subcloned and characterized. These fragments were also used as probes to isolate recombinant phage clones containing the cellular counterpart of the viral *src* gene, termed cellular *src*, from a lambda library of chicken deoxyribonucleic acid. The structure of cellular *src* was analyzed by restriction endonuclease mapping and electron microscopy. Restriction endonuclease mapping revealed extensive similarity between the *src* regions of Rous sarcoma virus and recovered avian sarcoma virus, but striking differences between the viral *src*'s and cellular *src*. Electron microscopic analysis of heteroduplexes between recovered virus *src* and cellular *src* revealed a 1.8-kilobase region of homology. In the cellular gene, the homologous region was interrupted by seven nonhomologous regions which we interpret to be intervening sequences. We estimate the minimum length of cellular *src* to be about 7.2 kilobases. These findings have implications concerning the mechanism of formation of recovered virus *src* and possibly other cell-derived retrovirus transforming genes.

It is now generally accepted that transforming genes of retroviruses are derived from loci in the genomes of vertebrate animals (4). *src* is the transforming gene of Rous sarcoma virus (RSV) and encodes a protein of about 60,000 daltons ($p60^{src}$) (9, 40). This protein has a protein kinase activity which specifically phosphorylates tyrosine residues in substrate proteins (12, 25, 32). The existence of *src*-related sequences in the genome of normal uninfected chicken cells was first demonstrated by Stehelin et al. (49). These *src*-related sequences are present in 26S polyadenylated messenger ribonucleic acid (RNA) molecules in uninfected chicken cells (23, 45, 51), although the extent of expression is limited (46, 56). The product from the endogenous sequences has been detected and identified as a 60,000-dalton protein, which is very similar to $p60^{src}$ in its structure and enzymatic activities (13, 29, 38). Therefore, this genetic element is called cellular *src* (c-*src*).

Hanafusa et al. (21) and Vigne et al. (52) reported the recovery of transforming viruses after infection of chickens with transformation-defective mutants of the Schmidt-Ruppin strain of RSV (SR-A) which contain less than 30% of the *src* sequences in their genome. Several lines of evidence showed that these viruses, named recovered avian sarcoma viruses (rASV), regained their transforming activity by acquiring intact *src* sequences through recombination between transformation-defective viruses and c-*src* (28, 29, 54, 55).

To gain information on the relationship between the *src* region of RSV (v-*src*) and c-*src* and clues to the origins of rASV, we isolated molecular clones containing *src* related sequences from the parental RSV, from rASV, and from c-*src*, the presumed progenitor of rASV. The following is a report of our initial analyses and characterization. Our studies show that the chromosomal c-*src* region has a more complex structure than the viral *src* sequences, resembling an "interrupted gene" of coding sequences separated by intervening sequences. These data are discussed in terms of possible mechanisms of recombination between c-*src* and transformation-defective RSV to generate the rASV.

MATERIALS AND METHODS

Cells and viral stocks. *Escherichia coli* ED8654 (37) was used for propagation of λ recombinant phages, and *E. coli* HB101 (8) was used for transformation and amplification of plasmids. Lambda vectors, λgtWES·B (31) and Charon 21A (5), were used for cloning of viral deoxyribonucleic acid (DNA). A lambda library of chicken DNA constructed by Dodgson et al. (17) was obtained from R. Axel, Columbia University, New York, N.Y. Briefly, this library was constructed as follows. After partial digestion of chicken DNA with *Hae*III and *Alu*I, fragments (15 to 20 kilobases [kb]) were purified by sedimentation in sucrose density gradients and ligated to λ Charon 4A arms by using *Eco*RI linkers.

Chicken embryo fibroblasts (CEF) were prepared from eggs (gs^- chf^-) obtained from SPAFAS, Inc. (Norwich, Conn.) and cultured as previously described (20). The rASV no. 1441 (rASV1441) used in this study was derived from a transformation-defective mutant, *td*108, of SR-A. The isolation and characterization of rASV1441 have been described (19, 21, 29, 54). Other avian retroviruses used were Rous-associated virus 2 (RAV2) and SR-A.

Molecular cloning. CEF were infected with either SR-A or rASV1441. Unintegrated circular DNA was prepared as a Hirt supernatant as previously described (26). Recombinant λ clones were isolated after partial digestion of viral DNA with *Eco*RI and ligation into the vector λgtWES·B. Procedures for packaging, sceening, purification, amplification, and preparation of DNA have been described (26). The preparation of the λ clone of RAV2 (λRAV2-1) has also been described (26). Lambda clones containing c-*src* were identified by using probes containing sequences from rv-*src* (see below). Approximately 300,000 plaque-forming units of phage were plated on 14 petri dishes (150 mm); these were screened by the procedure of Benton and Davis (2). Before hybridization filters were prewashed at 37°C in 5× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.02 M tris(hydroxymethyl)aminomethane-hydrochloride (pH 7.4) and 0.5% sodium dodecyl sulfate for 30 min. Hybridization was carried out at 37°C for 12 to 18 h in the above solution containing 50% formamide and an appropriate ^{32}P-labeled probe. After hybridization, the filters were washed twice with 0.2× SSC-0.2% sodium dodecyl sulfate solution and then incubated for 30 min in the same solution at 37°C. The filters were then rinsed with 0.2× SSC twice, dried, and exposed to Kodak XAR-5 film for 4 to 12 h with an intensifying screen. Positive plaques were purified four times. DNA from λ clones was prepared as previously described (57).

Subcloning of DNA fragments from cloned viral DNA in pBR322 was carried out by standard methods (6). The following modifications were used for the cloning of viral *src* containing *Eco*RI DNA fragments in the *Pst*I site of pBR322 to make selective use of loss of the ampicillin-resistant marker. The *Eco*RI fragments were first treated with DNA polymerase I in the presence of deoxyadenosine triphosphate and deoxyribosylthymine triphosphate to fill up the staggered structure at both ends (30). For this reaction we used 5 U of DNA polymerase I and 0.8 mM each deoxyadenosine triphosphate and deoxyribosylthymine triphosphate. Incubation was for 30 min at 20°C. After the repair reaction, residual substrates were removed by passing the mixture through a Sephadex G-75 column. The fragment was then incubated with terminal deoxynucleotidyl transferase in the presence of deoxycytidine triphosphate to add polydeoxycytidylic acid tails and annealed with pBR322 DNA which had been cleaved at the *Pst*I site and tailed with polydeoxyguanidylic acid (53). This modification made it possible to conserve *Eco*RI sites at both ends of the insert so that the insert could be excised from the recombinant molecule with *Eco*RI.

Transformation of *E. coli* was done as described by Villa-Komaroff et al. (53). Plasmid DNA was purified by the method of Clewell and Helinski (10).

Preparation of ^{32}P-labeled DNA probes. After restriction endonuclease digestion, appropriate DNA fragments were purified either by agarose gel electrophoresis or polyacrylamide gel electrophoresis. For agarose gel electrophoresis, low-melting-point agarose (Seaplaque agarose; Marine Colloids, Rockland, Maine) was used, and DNA fragments were purified according to a method developed by S. Astrin (personal communication). A gel section containing DNA fragments was transferred to an Eppendorf tube and heated at 65°C for 30 min and 37°C for 15 min. The melted solution was extracted twice with prewarmed (37°C), water-saturated phenol. NaCl was added to the aqueous phase to a final concentration of 0.1 M. Residual phenol was removed by centrifugation, and DNA was precipitated with ethanol. The DNA precipitate was washed with absolute ethanol three times and dried. Recovery of DNA from polyacrylamide gels was by electroelution (36). Purified DNA fragments were labeled by nick translation (33), yielding probes with specific activities of about 10^8 cpm/μg.

Restriction endonuclease mapping. Restriction endonucleases were purchased from Bethesda Research Laboratories (Bethesda, Md.) and New England Biolabs (Beverly, Mass.) and were used according to suppliers' instructions. DNA fragments were analyzed by gel electrophoresis, generally 0.8% agarose containing ethidium bromide (0.5 μg/ml) in tris(hydroxymethyl)aminomethane-acetate buffer (pH 7.8). After electrophoresis, DNA bands were visualized and photographed under ultraviolet light. DNA fragments which contained the *src* sequence were localized and mapped by Southern blotting (44) and hybridization with probes described below. The buffers used for hybridization and washing were the same as used for the screening of lambda clones as described above. λ DNA digested with *Eco*RI or *Hin*dIII were used as size markers.

Heteroduplex analysis. Heteroduplexes were formed in 50% formamide at an ionic strength (I) of 0.030 to 0.035 [Na^+ and tris(hydroxymethyl)aminomethane H^+] at 37°C. Spreading was performed by standard procedures (15) from a solution containing 55% formamide, 0.1 M tris(hydroxymethyl)aminomethane-hydrochloride (pH 8.5), 0.01 M ethylenediaminetetraacetate ($I = 0.06$), and 30 μg of cytochrome *c* per ml onto a water hypophase by a microspread

method (R.J., unpublished data). Phage ϕX174 single-strand and double-strand DNAs (5,386 base pairs) (41) were included in the spreads as size standards.

RESULTS

Molecular cloning of viral DNAs. Recombinant λ clones were prepared from unintegrated circular DNA of both SR-A and rASV1441. Restriction endonuclease mapping was used to determine the structure of the insert in the recombinant clones. Among the clones obtained, one representative clone was chosen from each group, and these clones, λSRA-V and λrASV-VIIB, were used for further analysis. *Eco*RI digestion of λSRA-V generated three large fragments (2.5, 2.0, and 1.5 megadaltons [Md]) and one small fragment (0.2 Md) from the insert (Fig. 1A). The presence of these three large fragments was consistent with the results reported by others (43, 50) and confirmed that this insert contains the entire genome of SR-A. The 0.2-Md fragment appeared to be derived from the long terminal repeat (LTR) region, since this fragment hybridized with a recombinant molecule containing only the LTR (27). These and other data provided the map shown in Fig. 1A. *Eco*RI digestion of λrASV-VIIB also gave three large fragments from the insert, but not the small (0.2-Md) fragment. No detectable difference was found in the molecular weights of the three fragments as compared with those of λSRA-V by gel electrophoresis. The results indicate that no significant deletion or addition is present in the genome of rASV, which is consistent with previous analyses of viral genome RNA by sedimentation in sucrose gradients (19) and by ribonuclease T_1-resistant oligonucleotide fingerprinting (54). Neither *Bam*HI digestion nor double digestion with *Bam*HI and *Eco*RI showed any difference between the two clones (data not shown), suggesting that no significant gross rearrangement had occurred. To further analyze the *src* region, the 2.0-Md *src*-containing fragments of SR-A and rASV1441 were subcloned in pBR322 as described below.

Subcloning of DNA fragments. Two fragments (2.7 and 2.4 Md) derived from λRAV2-1 after double digestion with *Hin*dIII and *Sal*I (7) were subcloned in pBR322 separately, yielding two subclones, pTT101 (2.7 Md) and pTT102 (2.4 Md) (Fig. 1B). These were used for detection of viral sequences in experiments described below.

The 2.0-Md *Eco*RI fragments of both λSRA-V and λrASV-VIIB, containing a part of *env*, the entire *src*, and *c*, were subcloned in the *Pst*I site of pBR322 without loss of the *Eco*RI site as described above. The subclones were designated pTT107 and pTT108, respectively (Fig. 1A). The structures of the *Eco*RI inserts from both plasmids were compared by restriction endonuclease mapping. For convenience, these inserts will be referred to as vs-RI and rvs-RI, respectively.

Some of the restriction endonuclease sites on rvs-RI are shown in Fig. 2A. The number and location of the restriction sites on vs-RI were the same as those of rvs-RI, except that the second *Pvu*II site in *c* was not present in vs-RI. Analysis of the nucleotide sequence in this region indicates that the absence of the *Pvu*II site in vs-RI is the result of a single base change (T.T., unpublished data). A significant difference was found between vs-RI and rvs-RI in the size of the smallest *Hae*II fragment which contains the intercistronic region between *env* and *src* (14; T.T., unpublished data): this fragment from vs-RI was about 15 base pairs longer than the corresponding fragment from rvs-RI. Except for these minor differences, the restriction endonuclease maps of vs-RI and rvs-RI were quite similar.

Preparation of specific DNA probes. The entire nucleotide sequence of vs-RI has been reported (14). From these results and our sequence data on rv-*src* (T.T., unpublished data) we have deduced the boundaries of the *env*, *src*, and *c* regions in rvs-RI. We made use of the restriction mapping data to isolate specific hybridization probes.

Six different probes prepared from rvs-RI were used for the analysis of c-*src*. The probes were prepared from DNA fragments purified and labeled as described above. Figure 2B illustrates the molecular origin of each probe: probe 1, the entire rvs-RI fragment, contains all of rv-*src*; probe 2, an *Ava*I-*Hin*cII fragment, includes the 5′ region of rv-*src*; probe 3, a *Bgl*I-*Eco*RI fragment, includes the 3′ region of rv-*src*; probe 4, the right-end *Pvu*II-*Eco*RI fragment, includes the 3′ end of rv-*src*; probe 5, the right-end *Hae*III-*Eco*RI fragment (because of the presence of many restriction sites for *Hae*III, its map is not shown in Fig. 2A), includes the repeated sequences located outside of both 5′ and 3′ termini of v-*src* as reported by Czernilofsky et al. (14); and probe 6, a fragment between the third and fourth *Ava*I sites, includes only rv-*src* sequences. pTT101 and pTT102 were combined and labeled as described above and used as the RAV2 probe.

Isolation of c-*src* recombinant phages. Approximately 300,000 plaque-forming units from the λ chicken library were screened by hybridization with probe 6. Three plaques which showed homology with this probe were isolated

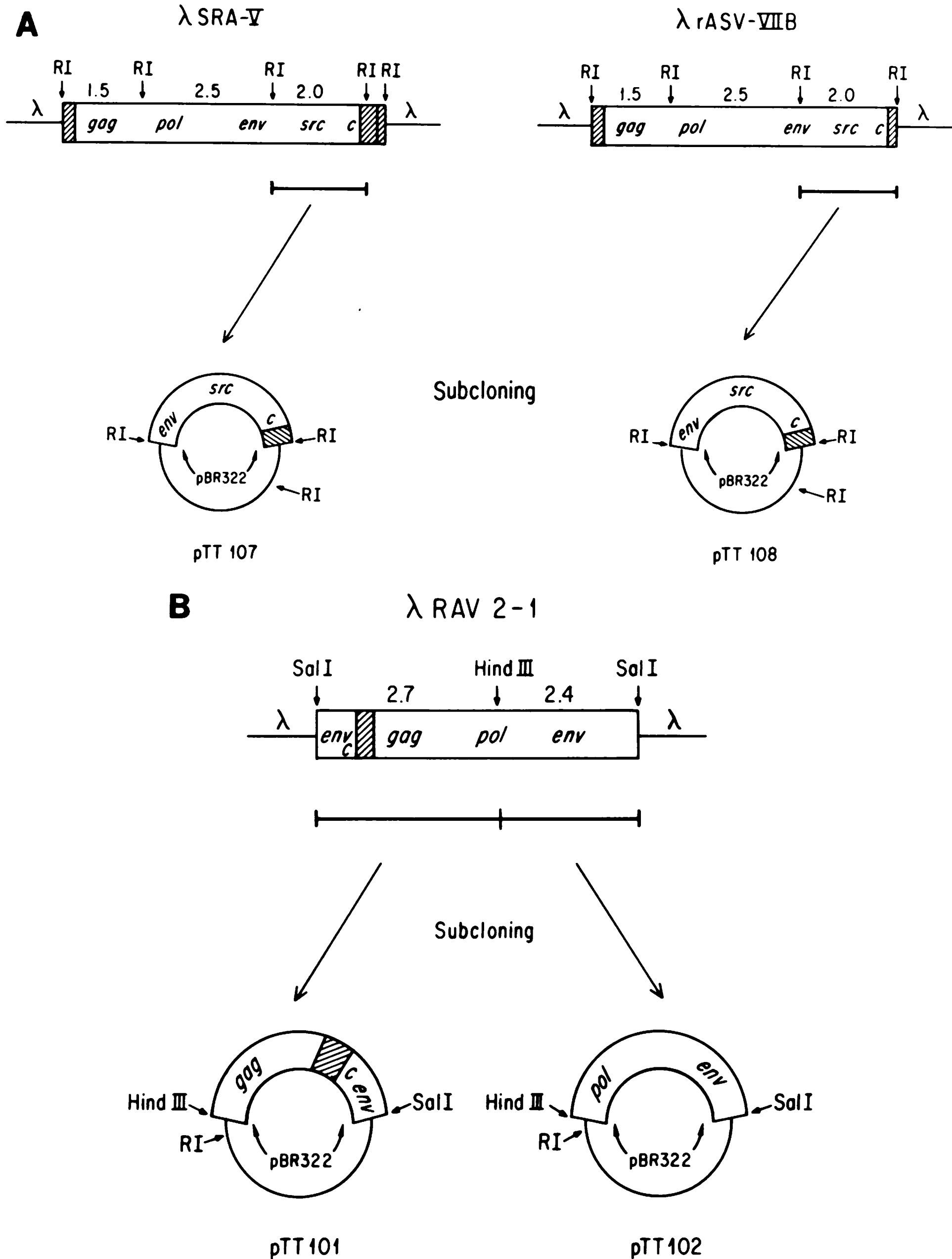


FIG. 1. *Molecular cloning of proviral DNA into lambda vectors and subcloning of derived DNA fragments into pBR322. Numbers between the EcoRI sites indicate the molecular mass (Md) of each fragment. The region shaded by diagonal lines is the LTR. RI, EcoRI. (A) Unintegrated circular proviral DNAs of SR-A and rASV1441 were partially digested with EcoRI and cloned in λgtWES·B. EcoRI sites present on each insert are indicated. λSRA-V contains two copies of the LTR, whereas λrASV-VIIB contains only one. The 2.0-Md EcoRI fragment containing src was subcloned in the PstI site of pBR322 as described in the text. pTT107 contains the viral src of λSRA, and pTT108 contains the src region of λrASV. (B) Unintegrated proviral DNA of RAV2 was digested with SalI and cloned in λCharon 21A (26). Two fragments (2.7 and 2.4 Md) generated by digestion with SalI and HindIII (7) were subcloned in pBR322. The recombinant plasmids pTT101 and pTT102 are shown below λRAV2-1.*

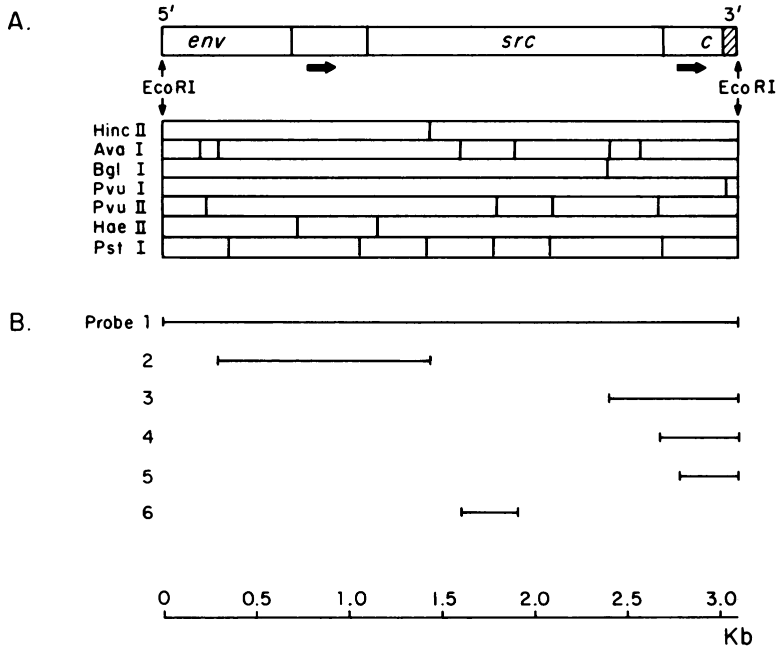


FIG. 2. *Restriction endonuclease map of the rvs-RI fragment from rASV1441 and derivation of specific probes. (A) Approximate boundaries of the env, src, and c regions are indicated in the boxed region which represents rvs-RI. The shaded region is a part of the LTR. These boundaries were deduced from the nucleotide sequence (14; T.T., unpublished data). Sites for seven restriction endonucleases are shown below. No digestion by the restriction enzymes BamHI, SacI, and HindIII was noted with this fragment. The horizontal arrows indicate the repeated sequences which flank the src gene (14). (B) Length and location of the DNA fragments used as specific probes for hybridization are indicated by solid lines. Probes: 1, entire rvs-RI fragment; 2, AvaI-HincII fragment; 3, BglI-EcoRI fragment; 4, PvuII-EcoRI fragment; 5, HaeIII-EcoRI fragment; 6, AvaI fragment.*

This is close to the expected number from the size of the haploid genome of chicken (1.4×10^6 kb) and the average size of insert DNA (15 to 20 kb). The three clones were designated λRCS3, λRCS14, and λRCS15, respectively. None of them showed homology with ^{32}P-labeled RAV2-derived plasmids pTT101 and pTT102 (Fig. 1). DNA isolated from these clones was analyzed by using restriction endonuclease digestions. The size of the insert in each clone was measured after digestion with *Eco*RI. Since the lambda vector used in preparation of the library contains no internal *Eco*RI sites, and *Eco*RI linkers were used in construction of the library, the sum of the lengths of *Eco*RI internal fragments in each clone equals the length of the chicken DNA insert. The sizes of the inserts were 18.4, 18.1, and 16.5 kb for λRCS3, λRCS14, and λRCS15, respectively (data not shown).

From our mapping data (Fig. 2A), *Ava*I digestion of rvs-RI generates three fragments (0.5, 0.3, and 0.18 kb), and *Pvu*II digestion generates two fragments (0.6 and 0.3 kb) from the internal region of rv-*src*. We digested DNA from the three c-*src* clones with these enzymes to determine whether fragments common to both rv-*src* and c-*src* were present. λRCS3 and λRCS14 gave patterns identical to each other: *Ava*I digestion yielded five bands and *Pvu*II yielded six bands which bound probe 1. However, none of these bands was identical in size to those of vs-RI (data not shown). These results indicate that the DNA structure of c-*src* is different from that of rv-*src*. λRCS15 generated only one band which hybridized with probe 1 after digestion with either *Pvu*II or *Ava*I. Thus, we conclude that λRCS15 contains only a part of the c-*src* region. Therefore, further study has focused on DNA from clones λRCS3 and λRCS14.

Patterns for restriction fragments of the two c-*src* clones analyzed by agarose gel electrophoresis and Southern blotting are shown in Fig. 3. For example, *Eco*RI digestion of the c-*src* clones gave four bands detectable by staining (Fig. 3A). Two of them (11 and 19.5 kb) were equal to the length of Charon 4A arms (16); the other two,

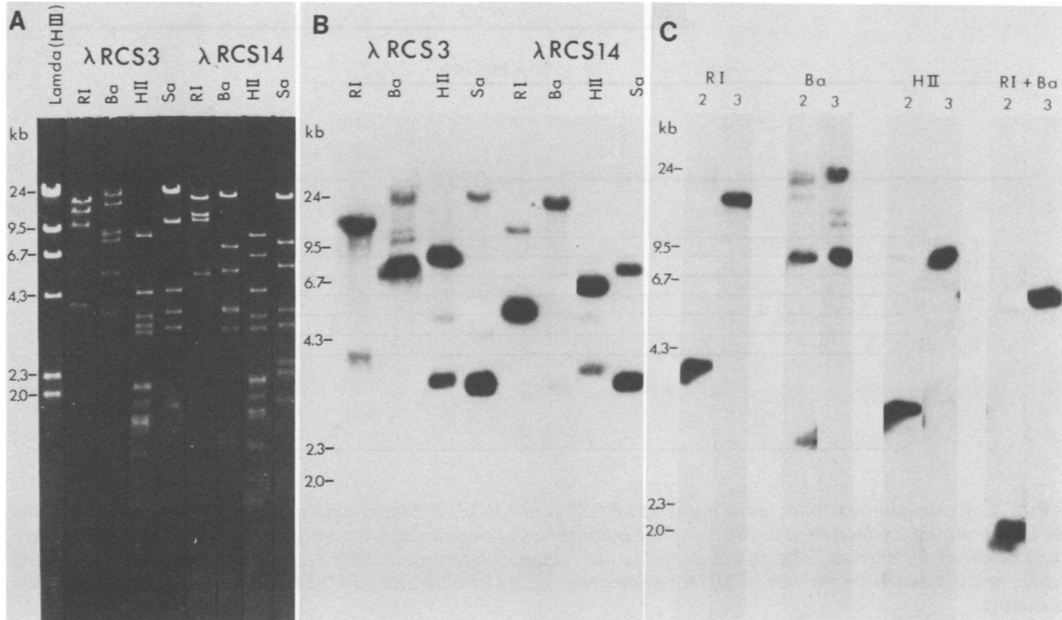


FIG. 3. *Restriction analyses and hybridization tests of fragments from c-src-containing λ clones. These analyses were done as described in the text. Restriction endonucleases are abbreviated as follows: Ba, BamHI; HII, HincII; HIII, HindIII; RI, EcoRI, and Sa, SacI. (A) DNAs from λRCS3 and λRCS14 were digested with restriction endonucleases as indicated, and fragments were visualized by staining. (B) DNAs from (A) were tested for homology with probe 1 (the entire rvs-RI). (C) DNA fragments from λRCS3 were tested for homology with probes 2 (5′ region) and 3 (3′ region) which are indicated by numbers above the lanes.*

presumed to be derived from the insert, hybridized with probe 1 (Fig. 3B). The extent of hybridization of the two fragments was unequal, however, suggesting that *Eco*RI cut asymmetrically within the gene. To determine polarity of the sites, DNA fragments were hybridized with probe 2 (for the 5′ region) and probe 3 (for the 3′ region) (Fig. 3C). For example, *Eco*RI generated two bands (14.4 and 4.0 kb) from λRCS3 which bound probe 1 (Fig. 3B, lane 1). Of these two fragments, the 4.0-kb fragment bound only probe 2 (Fig. 3C, lane 1), whereas the 14.4-kb fragment bound only probe 3 (Fig. 3C, lane 2). We concluded that the 4.0-kb fragment was derived from the 5′ region of c-*src* and the 14.4-kb fragment was derived from the 3′ region of c-*src*. Further analyses with other enzymes have been based on this unique *Eco*RI site which is therefore labeled "0" on the kb scale in Fig. 4. Nine different restriction endonucleases were used for this purpose. Except for *Hin*cII, which had a unique site on both fragments (Fig. 2A), none of the enzymes had restriction sites in either vs-RI or rvs-RI. The restriction sites mapped for some of the enzymes using procedures similar to those as described above for *Eco*RI are summarized in Fig. 4. Three enzymes (*Bcl*I, *Kpn*I, and *Sal*I) had no restriction sites within c-*src*.

Test for the presence of a 120-base-pair direct repeat in c-*src*. Czernilofsky et al. (14) reported the presence of repeated sequences (about 120 base pairs) which flank the 5′ and 3′ termini of the v-*src* from a SR-A clone. These sequences were also found in cloned DNA of rASV1441 (T.T., unpublished data). It was interesting to determine whether the c-*src* region also contained these sequences. Therefore, we hybridized probes 4 and 5 which contain these repeated sequences plus *c* and part of LTR (Fig. 5). As a control, we included RAV2 DNA, which is known to contain the sequences of *c* and part of LTR at 3′ end of the *env* gene (27). As expected, both probes 4 and 5 hybridized with RAV2 DNA. We also used *Hae*II digested rvs-RI to see whether the repeated sequence in *c* could hybridize with the repeated sequence which was located at the intercistronic region between *env* and *src* (14). *Hae*II digestion of vs-RI generates three fragments: 2.0 kb (containing *src* and *c*), 0.7 kb (containing *env*), and 0.4 kb (intercistronic region) (Fig. 2A). Probe 5 hybridized with the 2.0- and 0.4-kb fragments, but not with the 0.7-kb fragment (Fig. 5). These results showed that our test could detect homology in 90 out of 120 base pairs, the degree of similarity present in the direct repeat at the two locations

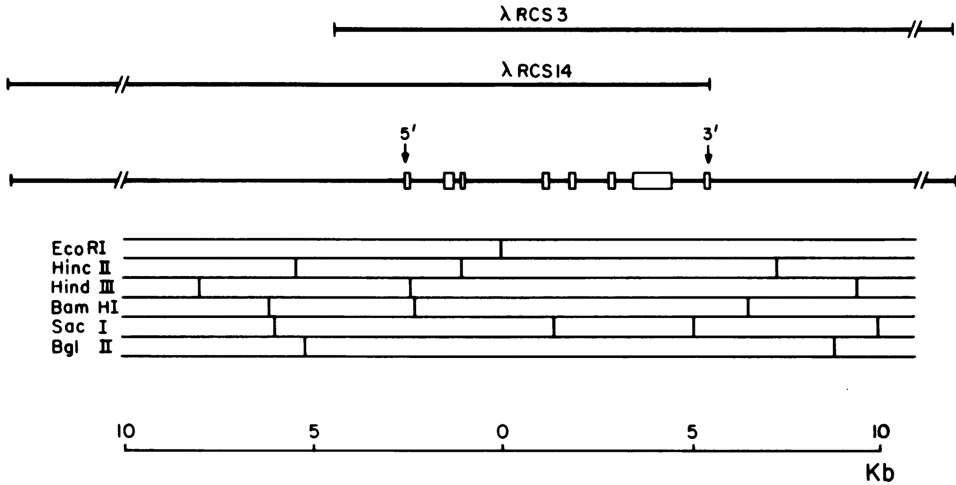


FIG. 4. *Restriction endonuclease maps of λRCS3 and λRCS14. The restriction sites of indicated enzymes were determined as described in the text. The boundaries of regions in c-src which are homologous with rv-src are indicated by arrows. The rv-src-homologous regions, determined by electron microscopy as described below, are indicated by the boxes. DNA regions carried by λRCS3 and λRCS14 are shown by the solid lines at the top.*

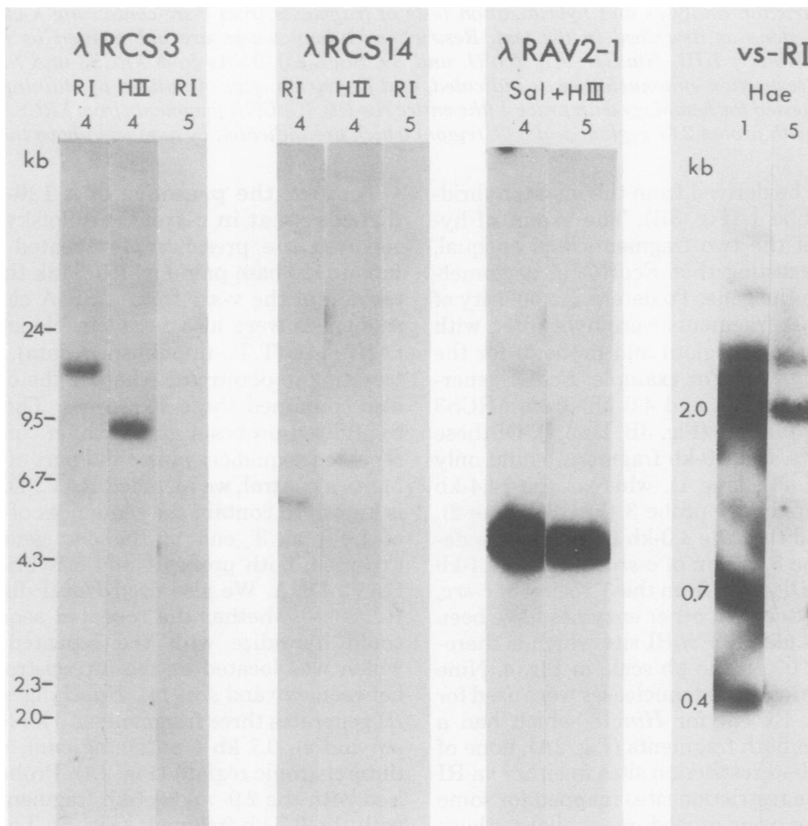


FIG. 5. *Tests for the presence of directly repeated sequences in λRCS3 and λRCS14. Restriction endonucleases are abbreviated as in Fig. 3, except Sal indicates SalI and Hae indicates HaeII. One microgram of each DNA was added to each lane, and probes 1, 4, and 5 were used as indicated above the lanes.*

(14). When these probes 4 and 5 were tested on λRCS3 and λRCS14, probe 4 hybridized with either *Eco*RI or *Hin*cII fragments which derived from the 3′ region of λRCS3 and λRCS14, whereas no hybridization was observed between the probe 5 and either λRCS3 or λRCS14. Thus, we conclude that the repeated sequences present in the viral genome are not derived from cellular sequences included in the c-*src* regions we have cloned.

Authenticity of c-*src* region in recombinant clones. Next, we compared the structure of cloned c-*src* with that in chicken genome DNA. For this purpose we extracted DNA from the erythrocytes of a chicken with no endogenous virus sequences (1). The results after digestion of the cellular DNA with *Bam*HI and double digestion with *Bam*HI and *Eco*RI and Southern blotting are shown in Fig. 6. Two fragments (8.7 and 4.0 kb) hybridized with the probe 1 and 2 after *Bam*HI digestion. Additional digestion by *Eco*RI cleaved the larger fragment (8.7 kb) into two fragments (6.5 and 2.1 kb), of which only the smaller fragment (2.1 kb) hybridized with the probe 2. These and additional data with other enzymes gave results identical to those obtained from restriction analysis of

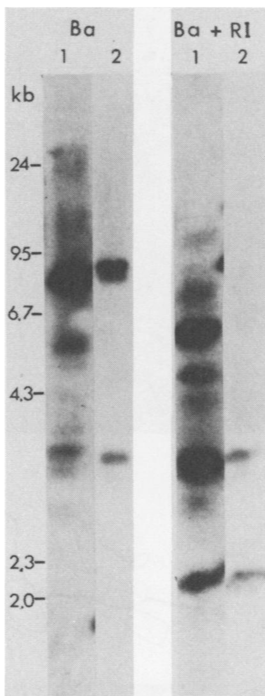


FIG. 6. *Gel electrophoresis and hybridization analysis of restriction fragments of chicken genomic DNA. Probes and enzyme designations are as in Fig. 3.*

cloned DNAs shown in Fig. 4. Thus, we conclude that the DNA in our clones is a faithful representation of the *src* region in the chicken genome.

Furthermore, the fact that single bands which contained the boundary region of c-*src* and either 5′ or 3′ flanking sequences (e.g., Fig. 6, second and fourth lanes) were observed after digestion with several restriction enzymes strongly suggests that a single copy of c-*src* exists in genome DNA, as proposed by McClements et al. (34).

Electron microscope studies. Heteroduplex comparison between the two clones λRCS3 and λRCS14 showed that they contain inserts in opposite orientations. Figure 7A shows no hybridization between the inserts when the lambda arms are annealed. Conversely, as shown in Fig. 7B, when the inserts are hybridized, the lambda arms are not annealed. The extent of overlap between the inserts in the clones (Fig. 4), determined from the double strand region of examples of the type shown in Fig. 7B, was 8.37 ± 0.67 kb ($n = 5$).

Studies were then undertaken to compare the c-*src* clones with the rv-*src* sequence in rvs-RI. Since the data described below indicated that λRCS3 contains more of the c-*src* region than λRCS14, our analyses focused on λRCS3. Examples of such heteroduplexes are shown in Fig. 7 (C and D).

The beginning of the 5′ end of the rv-*src* homologous sequences in λRCS3 was located approximately 1.9 ± 0.7 kb ($n = 6$) from the left arm of lambda, and the 3′ end was located about 10.1 ± 0.8 kb ($n = 5$) from the right arm of lambda by using 19.8 and 10.9 kb as the Charon 4A RI-arm lengths (16). With the lower limit in length of the c-*src* determined to be 7,210 ± 360 nucleotides (see below), the sum of the insert lengths yields a figure of 19.2 ± 1.0 kb, comparable to the 18.4-kb insert length determined by gel electrophoresis of restriction fragments (above). In all, seven deletion loops and eight homology regions were detected. The smallest of the deletion loops, indicated by a black ball on the diagram, is difficult to visualize in the examples shown in Fig. 7C and D. Nevertheless, analyses of many molecules verified its presence (Table 1, no. 4). The polarity of the hybrid was established on the basis of the lengths of rvs-RI sequence protruding as single-strand tails from either end of the hybrid region. Length determinations are summarized in Table 1.

The terminal loops (no. 2 and 14) were not present in every heteroduplex molecule. The 5′ loop (no. 2) was present in roughly 80% (37 of 47) of the molecules examined, whereas the 3′ loop (no. 14) was present only half as frequently

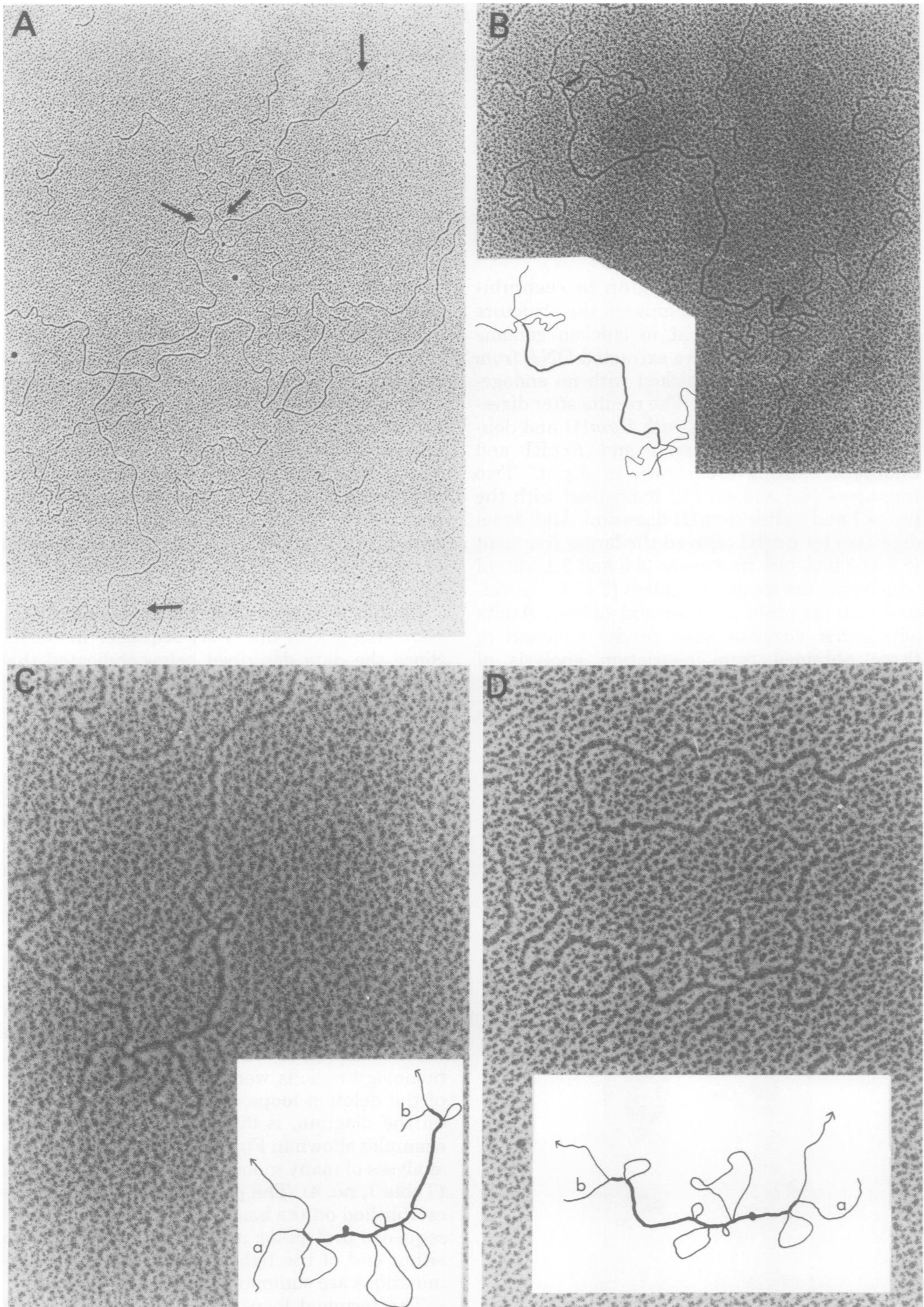
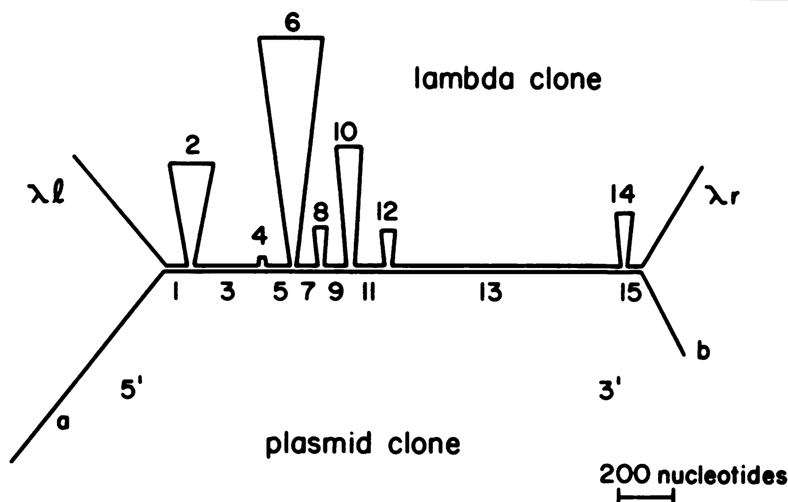


FIG. 7. ***Analysis by electron microscopy. (A and B) Heteroduplexes between λRCS3 and λRCS14. (A) Duplex regions delimited by arrows (19.8 and 10.9 kb) are from lambda arms. Internal single-strand loops are the clone inserts. (B) Duplex region (8.3 ± 0.7 kb) is from insert containing the c-src region and 5′ flanking sequences (see text). Single-stranded regions are the lambda arms. These results show the polarity of inserts is different between the two clones. (C and D) Heteroduplexes between λRCS3 and rvs-RI fragment. a, Unhybridized 5′ region of the rvs-RI fragment. b, Unhybridized 3′ region of the rvs-RI fragment. Odd numbers indicate the duplex regions, and even numbers indicate the loops. The numbering starts from 5′ end (Table 1).***

TABLE 1. *Summary of length measurements by electron microscopy*[a]



Features		Measurements				
		Single-stranded		Double-stranded		
Loop	Duplex	Raw Measure	Corrected[b]	Raw Measure	Corrected[b]	Number
	1			102 ± 26	97 ± 25	27
2		1050 ± 136[c]	997 ± 129			30
	3			259 ± 27	246 ± 26	28
4		49 ± 21	47 ± 20			37
	5			123 ± 16	119 ± 15	34
6		2146 ± 185	2040 ± 176			42
	7			120 ± 29	114 ± 28	47
8		411 ± 62	390 ± 59			36
	9			122 ± 28	116 ± 27	44
10		1065 ± 124	1012 ± 118			46
	11			153 ± 16	145 ± 15	38
12		371 ± 43	352 ± 41			33
	13			937 ± 74	890 ± 70	31
14		593 ± 102	563 ± 97			17
	15			89 ± 22	84 ± 21	15
TOTALS		5685 ± 290[d]	5400 ± 270	1905 ± 100	1810 ± 90	
Tail						
a		991 ± 83	940 ± 79			24
b		364 + 40	346 ± 38			11

[a] Measurements were derived relative to ϕX double-stranded DNA ($n = 9$; coefficient of variance = 4.4%) and single-stranded DNA ($n = 12$; coefficient of variance = 4.6%) standards of 5,386 nucleotides. A double-strand/single-strand ratio of 1.035 was observed. Loop and duplex features are numbered according to the accompanying diagram. λl and λr indicate left and right arms of lambda DNA. Sequences a and b are tails corresponding to non-*src* sequences in the rvs-RI fragment which do not anneal (Fig. 7).

[b] The sum of the raw measure double-stranded and the rvs-RI fragment tails, a and b, equals 3,260. This length by sequencing should be 3,110 (T. T., unpublished data). To correct for over-measurement, a factor of 0.95 is applied.

[c] Mean ± unbiased standard deviation.

[d] The standard deviation of the sum is derived as the square root of the sum of the variances. The values are approximate because covariances are assumed to be zero, and minor corrections to compensate for sample size differences are neglected.

(19 of 47). The terminal duplex regions (no. 1 and 15) are of the order of 50 to 100 nucleotides in length by microscopy and may be of limited stability in the formamide solvents employed in the hybridization and spreading. A cursory examination after spreading in 45% formamide showed no obvious increase in the presence of the 3′ loop and suggests that the hybridization rather than the spreading provided the stringency. This observation is consistent with the 17°C difference in stringency (considering effects of temperature, salt, and formamide concentration [35, 42]) between the hybridization and the spreading conditions. Thus, the brief period during which hybrids are held at room temperature was insufficient to complete the annealing reaction.

The 3′ loop (no. 14) was not observed in λRCS14. This is surprising given that probe 4, containing 30 nucleotides at the 3′ end of the rv-*src*, hybridized to λRCS 14. This hybridization, however, appeared weaker than that with an equivalent mass of DNA from λRCS3 (Fig. 5, lanes 1, 2, 4, and 5). A plausible explanation is that the λRCS14 insert begins within the last duplex region (no. 15), leaving a length adequate to be detected by Southern hybridization but inadequate to form stable hybrids in the solvent conditions used for heteroduplex formation.

In this latter regard, it should be noted that small lengths such as that of the 3′ and 5′ extreme duplex regions (no. 1 and 15) and of the small loop (no. 4) can be easily overestimated. The width of cytochrome-coated, double-stranded DNA is about 15.0 nm and corresponds to a length of 50 nucleotides. It is unclear where the ends of structures lie within the cytochrome aggregate, and the particular choice of beginning and end points for a double-stranded feature during the tracing can influence the length by ±50 nucleotides. As the outside termini of duplexes no. 1 and 15 were measured by using the far edge of the cytochrome aggregates, the length values of approximately 100 and 80 nucleotides may be regarded as maximum estimates; they could possibly be as low as 50 and 30 nucleotides, respectively.

The length of the rvs-RI was measured as 3,110 nucleotides by sequencing (T.T., unpublished data). The measurement by electron microscopy gave a figure of 3,077 ± 40 nucleotides ($n = 10$) relative to phage ϕX174 standard (5,386 nucleotides). When the sum of the lengths of the plasmid tails and the duplex regions in the hybrid is taken, a length figure of 3,260 is obtained, for which statistical analysis yields a standard deviation of about 140. The difference between these two measurements is significant at the $\alpha < 0.001$ level by a two-tailed t test. The difference could arise from a tendency to over-measure lengths during tracing of multiple features adjacent within the same molecule. We do not have a similar control for the sum of nucleotides in the nonhomologous single-stranded regions and therefore do not know whether any bias arose in these measurements. To compensate for possible systematic errors in length measurement, a correction of 3,110/3,260 = 0.95 has been applied to all of the determinations.

With corrections applied, the rv-*src* in the plasmid (the sum of duplex regions) is 1,810 nucleotides, and an approximate analysis yields a standard deviation of 100. The corresponding figures for the sum of the deletion loops are 5,400 ± 270 nucleotides. The total length of the *src*-related sequences in lambda is 7,200 ± 340 nucleotides. None of the loops, except the smallest (no. 4), could be assigned to sequences from the plasmid without violating the plasmid length constraints at very high levels of significance ($\alpha < 0.001$).

Figure 4 and Table 1 summarize conclusions on the structural relationship between the viral and cellular *src* clones as derived from these electron microscope studies and restriction endonuclease mapping.

DISCUSSION

We have compared cloned DNAs containing sequences homologous to *src* from three sources: v-*src* (λSRA), rv-*src* (λrASV) and c-*src* (λRCS). Analysis of the DNA structure of the c-*src*-containing clones indicates that the c-*src* sequences we have cloned are representatives of the single copy of this region present in the chicken genome. (i) Restriction maps of the *src*-containing regions in λRCSs and normal cellular DNA are virtually identical. (ii) The formation of unique bands of DNA fragments containing sequences flanking c-*src* indicate that only one copy of c-*src* is present in the chicken genome. (iii) The lower limit in size of c-*src* gene based on homology with rv-*src* is about 7.20 kb, whereas the duplex region between λRCS3 and λRCS14 (Fig. 7A and B) is about 8.37 kb. This suggests that 5′ flanking sequences are homologous between the two independent c-*src* clones, and that the two clones of c-*src* are derived from the same locus in the chicken genome.

The restriction maps of vs-RI and rvs-RI are very similar. The only differences found are the presence of one new site in rvs-RI within the *src* region and the size difference in one restriction fragment located outside the 5′ end of *src*. The overall similarity of these two clones has been

recently verified by nucleotide sequence analysis (T.T., unpublished data). In contrast, the structure of c-*src* is strikingly different. Several enzymes which cannot digest rv-*src* have restriction sites in c-*src*. Some enzymes which cleave rv-*src* cleave c-*src* at increased numbers of sites. More significantly, all of the fragments generated by restriction enzymes from c-*src* were different in size from those derived from rv-*src*. Therefore, c-*src* must have a more complex structure than rv-*src*. This was verified by direct comparison of viral and cellular *src* DNAs by electron microscopy.

Electron microscopic analysis of the heteroduplexes shows that the heteroduplexes contain seven deletion loops and eight common regions. The combined total length of the common regions is about 1.8 kb, and the total length calculated for c-*src* is 7.2 kb. Since rv-*src* contains all of the information for production of the *src* protein ($p60^{src}$) (22, 29), its sequence must contain all of the structural elements of the rv-*src* gene. It seems likely, therefore, that the extra DNA in the c-*src* counterpart represents nontranslated "intervening sequences" whose transcript is spliced out during messenger RNA processing. The total length of the eight coding sequences is equivalent to the size of rv-*src*. The nearly complete duplex formation is consistent with the extensive cross-hybridization between complementary DNA against v-*src* and normal cellular RNA (47, 56).

Electron microscopic measurements provide a lower limit for the size of the c-*src* gene of 7.2 kb. We do not know the maximal length of the c-*src* gene, since it is possible that some cellular coding regions which are not homologous to rv-*src* exist at both the 3′ and 5′ ends of c-*src*. Further, if other intervening sequences exist, the c-*src* gene could be considerably larger. However, because the size and primary structure of gene products from c-*src* and v-*src* (and rv-*src*) are almost identical (60 kilodaltons) (13, 28, 38), and because the length of the homologous region (duplex region) shared by c-*src* and rv-*src* is sufficient to code for a 60-kilodalton protein, the extra coding sequences, if needed, must be very small. Previously, messenger RNA of c-*src* has been shown to sediment at about 26S in a sucrose gradient (23, 45, 56); this value is roughly equivalent to about 4.4 kb by Spirin's formula (48). The 4.4-kb c-*src* messenger RNA may contain information of a large leader sequence, 3′ nontranslated region or another cellular gene. Absence of homology with RAV2 indicates that no sequences similar to viral LTR lies within 12.6 kb of the 5′ end of the c-*src* coding region in the λRCS14 clone.

After these studies were completed, results from similar analyses were communicated to us by Cooper and his associates (42a), who compared the structures of v-*src* from SR-D with four separate clones containing different regions of c-*src* derived from the same λ library that we used. Their results are similar to ours, except that they did not observe the last loop (no. 14) and duplex region (no. 15) at the 3′ end of the gene. This discrepancy could be due to the instability of this loop as discussed before, or due to the small number of molecules they examined in which the pattern of the loops' appearance did not become evident. In any case, their estimate of the extent of the chicken genome which encompasses the *src* gene is accordingly less than ours.

It now appears that normal vertebrate cells contain in their genomes counterparts of most, if not all, of the transforming sequences of oncogenic retroviruses, and the structures of some of these cellular sequences have been analyzed (3, 18, 39). Evidence has also accumulated to indicate that these transforming viruses are derived by recombination of viral and host cell sequences (22). The rASV system is a useful model since the frequency of recombination is relatively high, and all interacting elements (i.e., transformation-defective mutants of SR-A, c-*src*, and rASV) can be analyzed.

Basically, two possibilities are conceivable for the mechanism of formation of rASV. One mechanism is homologous recombination at the DNA level. Normal cellular *src* might be inserted at and possibly replace the *src* region which is partially retained in the transformation-defective SR-A virus after recombination in the host chromosome at the c-*src* locus, producing a large recombinant containing the intervening sequences in c-*src*. The transcript of this recombinant would be spliced to make a smaller rASV genome which could be packaged into infectious virions. In this model, we would assume that the removal of the c-*src* intervening sequences occurs as a normal consequence of splicing of viral RNA. The alternative model considers that the recombination event occurs at the RNA level. A messenger RNA containing c-*src* would be incorporated into virions. Recombination could occur at the step of reverse transcription (11) or between proviral DNA intermediates (24). At the moment, we cannot determine which of these two alternatives operates in the formation of rASV. Analyses of DNA of primary tumors from which rASV's were isolated are in progress and could be helpful in distinguishing them.

Yamamoto et al. (58) first identified a short stretch of similar sequences distal to the 3′ end

of the *env* and *src* regions of RSV. The presence of these sequences as a direct repeat at both 5′ and 3′ flanking sequences of *src* was demonstrated by Czernilofsky et al. (14). These sequences are probably involved in the deletion of *src* from nondefective RSV by homologous recombination (14, 58). It has been suggested that these sequences may also function in insertion of *src* (58). However, since they do not appear to flank the region in c-*src* which is homologous to rv-*src*, our results seem to rule out a simple model whereby copies of the direct repeat in the cellular genome facilitate recombination.

ACKNOWLEDGMENTS

We thank R. Axel for providing the lambda library of chicken DNA.

This work was supported by Public Health Service grants CA 14935 and CA 18213 from the National Cancer Institute.

ADDENDUM

DNA sequencing of "duplex region no. 1" of cloned c-*src* (λRCS3) revealed that the homologous region consists of 81 nucleotides. This is very close to the size estimated from the heteroduplex analysis described in this paper. The sequencing also showed that duplex region no. 1 contains the nucleotide sequence equivalent ot −92 to −12 nucleotides in the intercistronic region of rvs-RI (+1 represents the first nucleotide in the translated frame of $p60^{scr}$).

LITERATURE CITED

1. **Astrin, S. M., E. G. Buss, and W. S. Hayward.** 1979. Endogenous viral genes are non-essential in the chicken. Nature (London) **282:**339–341.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196:**180–182.
3. **Beveren, C. V., J. A. Galleshaw, V. Joans, A. J. M. Berns, R. F. Doolittle, D. J. Donoghue, and I. M. Verma.** 1981. Nucleotide sequence and formation of the transforming gene of a mouse sarcoma virus. Nature (London) **289:**258–262.
4. **Bishop, J. M.** 1981. Enemies within: the genesis of retrovirus oncogenes. Cell **23:**5–6.
5. **Blattner, F. R., D. O. Kiefer, D. D. Moore, J. R. de Wet, and B. G. Williams.** 1978. Application for EK2 certification of a host-vector system for DNA cloning, supplement 9. Data on Charon 21A.
6. **Bolivar, F. R., L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2:**95–113.
7. **Boone, L. R., and A. M. Skalka.** 1981. Viral DNA synthesized *in vitro* by avian retrovirus particles permeabilized with melittin. II. Evidence for a strand displacement mechanism in plus-strand synthesis. J. Virol. **37:**117–126.
8. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41:**459–472.
9. **Brugge, J. S., and R. L. Erikson.** 1977. Identification of a transformation-specific antigen induced by an avian sarcoma virus. Nature (London) **269:**346–348.
10. **Clewell, D. B., and D. R. Helinski.** 1969. Supercoiled circular DNA-protein complex in Escherichia coli: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. U.S.A. **62:**1159–1166.
11. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genome: some unifying hypotheses. J. Gen. Virol. **42:**1–26.
12. **Collet, M. S., and R. L. Erikson.** 1978. Protein kinase activity associated with the avian sarcoma virus *src* gene product. Proc. Natl. Acad. Sci. U.S.A. **75:**2021–2024.
13. **Collett, M. S., J. S. Brugge, and R. L. Erikson.** 1978. Characterization of a normal avian cell protein related to the avian sarcoma virus transforming gene product. Cell **15:**1363–1369.
14. **Czernilofsky, A., A. Levinson, H. Varmus, J. M. Bishop, E. Tischer, and H. Goodman.** 1980. Nucleotide sequence of an avian sarcoma virus oncogene (*src*) and proposed amino acid sequence for gene product. Nature (London) **287:**198–203.
15. **Davis, R. W., M. Simon, and N. Davidson.** 1971. Electron microscope heteroduplex methods for mapping regions of base sequence homology in nucleic acids. Methods Enzymol. **21:**413–428.
16. **de Wet, J. R., D. L. Daniels, J. L. Schroeder, B. G. Williams, K. Denniston-Thompson, D. D. Moore, and F. R. Blattner.** 1980. Restriction maps of twenty-one Charon vector phages. J. Virol. **33:**401–410.
17. **Dodgson, J. B., J. Strommer, and J. D. Engel.** 1979. Isolation of the chicken β-globin gene and a linked embryonic β-like globin gene from a chicken DNA recombinant library. Cell **17:**879–887.
18. **Goff, S. P., E. Gilboa, O. N. Witte, and D. Baltimore.** 1980. Structure of the Abelson murine leukemia virus genome and the homologous cellular gene studies with cloned viral DNA. Cell **22:**777–785.
19. **Halpern, C. C., W. S. Hayward, and H. Hanafusa.** 1979. Characterization of some isolates of newly recovered avian sarcoma virus. J. Virol. **29:**91–101.
20. **Hanafusa, H.** 1969. Rapid transformation of cells by Rous sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **63:**318–325.
21. **Hanafusa, H., C. C. Halpern, D. L. Buchhagen, and S. Kawai.** 1977. Recovery of avian sarcoma virus from tumors induced by transformation-defective mutants. J. Exp. Med. **146:**1735–1747.
22. **Hanafusa, H., L.-H. Wang, T. Hanafusa, S. M. Anderson, R. E. Karess, and W. S. Hayward.** 1980. The nature and origin of the transforming gene of avian sarcoma viruses, p. 483–497. *In* B. N. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.
23. **Hayward, W. S.** 1977. Size and genetic content of viral RNAs in avian oncovirus infected cells. J. Virol. **24:**47–63.
24. **Hunter, E.** 1978. The mechanism for genetic recombination in the avian retroviruses. Curr. Top. Microbiol. Immunol. **79:**295–309.
25. **Hunter, T., and B. Sefton.** 1980. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc. Natl. Acad. Sci. U.S.A. **77:**1311–1315.
26. **Ju, G., L. Boone, and A. M. Skalka.** 1980. Isolation and characterization of recombinant DNA clones of avian retroviruses: size heterogeneity and instability of the direct repeat. J. Virol. **33:**1026–1033.
27. **Ju, G., and A. M. Skalka.** 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. Cell **22:**379–386.
28. **Karess, R. E., and H. Hanafusa.** 1981. Viral and cellular *src* genes contribute to the structure of recovered avian sarcoma virus transforming protein. Cell **24:**155–164.
29. **Karess, R. E., W. S. Hayward, and H. Hanafusa.** 1979. Cellular information in the genome of recovered avian

sarcoma virus directs the synthesis of transforming protein. Proc. Natl. Acad. Sci. U.S.A. **76**:3154–3158.

30. **Kleppe, K., E. Ohtsuka, R. Kleppe, I. Molineux, and H. G. Khorana.** 1971. Studies on polynucleotides. XCVI. Repair replication of short synthetic DNA's as catalyzed by DNA polymerases. J. Mol. Biol. **56**:341–361.
31. **Leder, P., D. Tiemeier, and L. Enquist.** 1977. EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the λgtWES system. Science **196**:175–177.
32. **Levinson, A. D., H. Oppermann, L. Levintow, H. E. Varmus, and J. M. Bishop.** 1978. Evidence that the transforming gene of avian sarcoma virus encodes a protein kinase associated with a phosphoprotein. Cell **15**:561–572.
33. **Maniatis, T., A. Jeffrey, and D. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage lambda. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
34. **McClements, W., H. Hanafusa, S. Tilghman, and A. M. Skalka.** 1979. Structural studies on oncornavirus related sequences in chicken genomic DNA: two-step analysis of *Eco*RI and *Bgl*I restriction digests and tentative mapping of a ubiquitous endogenous provirus. Proc. Natl. Acad. Sci. U.S.A. **76**:2165–2169.
35. **McConaughy, B., C. Laird, and B. McCarthy.** 1969. Nucleic acid reassociation in formamide. Biochemistry **8**:3289–3295.
36. **McDonell, M. W., M. N. Simon, and F. W. Studier.** 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gel. J. Mol. Biol. **110**:119–146.
37. **Murray, N. E., W. J. Brammer, and K. Murray.** 1977. Lambdoid phages that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. **150**:53–61.
38. **Oppermann, H., A. Levinson, H. E. Varmus, L. Levintow, and J. M. Bishop.** 1979. Uninfected vertebrate cells contain a protein that is closely related to the product of the avian sarcoma virus transforming gene (*src*). Proc. Natl. Acad. Sci. U.S.A. **76**:1804–1808.
39. **Oskarsson, M., W. L. McClements, D. G. Blair, J. V. Maizel, and G. F. Vande Woude.** 1980. Properties of a normal mouse cell DNA sequence (*sarc*) homologous to the *src* sequence of Moloney sarcoma virus. Science **207**:1222–1224.
40. **Purchio, A. F., E. Erikson, J. S. Brugge, and R. L. Erikson.** 1978. Identification of a polypeptide encoded by the avian sarcoma virus *src* gene. Proc. Natl. Acad. Sci. U.S.A. **75**:1567–1571.
41. **Sanger, F., G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III, P. M. Slocombe, and M. Smith.** 1977. Nucleotide sequence of bacterophage φX174 DNA. Nature (London) **265**:687–695.
42. **Schildkraut, C., and S. Lifson.** 1965. Dependence of the melting temperature of DNA on salt concentration. Biopolymers **3**:195–208.

42a. **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.

43. **Shank, P. R., S. H. Hughes, H.-J. Kung, J. E. Majors, N. Quintrell, R. V. Guntaka, J. M. Bishop, and H. E. Varmus.** 1978. Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell **15**:1383–1395.
44. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
45. **Spector, D. H., B. Baker, H. E. Varmus, and J. M. Bishop.** 1978. Characterization of cellular RNA related to the transforming gene of avian sarcoma viruses. Cell **13**:381–386.
46. **Spector, D. H., K. Smith, T. Padgett, P. McCombe, D. Roulland-Dussoix, C. Moscovici, H. E. Varmus, and J. M. Bishop.** 1978. Uninfected avian cells contain RNA related to the transforming gene of avian sarcoma viruses. Cell **13**:371–379.
47. **Spector, D. H., H. E. Varmus, and J. M. Bishop.** 1978. Nucleotide sequences related to the transforming gene of avian sarcoma virus are present in DNA of uninfected vertebrates. Proc. Natl. Acad. Sci. U.S.A. **75**:4102–4106.
48. **Spirin, A. S.** 1963. Some problems concerning macromolecular structure of RNA. Prog. Nucleic Acid Res. Mol. Biol. **1**:301–345.
49. **Stehelin, D., H. E. Varmus, J. M. Bishop, and P. K. Vogt.** 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature (London) **260**:170–173.
50. **Taylor, J. M., T. W. Hsu, and M. M. C. Lai.** 1978. Restriction enzyme sites on the avian RNA tumor virus genome. J. Virol. **26**:479–484.
51. **Varmus, H. E., D. Stehelin, D. Spector, J. Tal, F. Fujita, T. Padgett, D. Roulland-Dussoix, H.-J. Kung, and J. M. Bishop.** 1976. Distribution and function of defined regions of avian tumor virus genomes in viruses and uninfected cells, p. 339–358. *In* D. Baltimore, A. S. Huang, and C. F. Fox (ed.), Animal virology. Academic Press, Inc., New York.
52. **Vigne, R., M. L. Breitman, C. Moscovici, and P. K. Vogt.** 1979. Restitution of fibroblast-transforming ability in *src* deletion mutants of avian sarcoma virus during animal passage. Virology **93**:413–426.
53. **Villa-Komaroff, L., A. Efstratiadis, S. Broome, P. Lomedico, R. Tizard, S. P. Naber, W. L. Chick, and W. Gilbert.** 1978. A bacterial clone synthesizing proinsulin. Proc. Natl. Acad. Sci. U.S.A. **75**:3727–3731.
54. **Wang, L.-H., C. C. Halpern, M. Nadel, and H. Hanafusa.** 1978. Recombination between viral and cellular sequences generates transforming sarcoma viruses. Proc. Natl. Acad. Sci. U.S.A. **75**:5812–5816.
55. **Wang, L.-H., C. Moscovici, R. E. Karess, and H. Hanafusa.** 1979. Analysis of the *src* gene of sarcoma viruses generated by recombination between transformation-defective mutants and quail cellular sequences. J. Virol. **32**:546–556.
56. **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.
57. **Williams, B. G., F. R. Blattner, S. R. Jaskunas, and M. Nomura.** 1977. Insertion of DNA carrying ribosomal protein genes of *Escherichia coli* into Charon vector phages. J. Biol. Chem. **252**:7344–7354.
58. **Yamamoto, T., J. S. Tyagi, J. B. Fagan, G. Jay, B. deCrombrugghe, and I. Pastan.** 1980. Molecular mechanism for the capture and excision of the transforming gene of avian sarcoma virus as suggested by analysis of recombinant clones. J. Virol. **35**:436–443.