MOLECULAR AND CELLULAR BIOLOGY, Sept. 1984, p. 1738–1746
0270-7306/84/091738-09$02.00/0

Vol. 4, No. 9

Mutation of a Termination Codon Affects *src* Initiation

STEPHEN HUGHES,[1]†* KARIN MELLSTROM,[2] ELAINE KOSIK,[1]† FUYUHIKO TAMANOI,[1] AND JOAN BRUGGE[2]

Cold Spring Harbor Laboratory, Cold Spring Harbor,[1] and Department of Microbiology, State University of New York at Stony Brook, Stony Brook,[2] New York 11724

Received 15 March 1984/Accepted 14 June 1984

The four Rous sarcoma virus messages *gag*, *gag-pol*, *env*, and *src* all derive from a full-length RNA precursor. All four messages contain the same 5′ leader segment. Three of the messages, *gag*, *gag-pol*, and *env*, use an AUG present in this leader to initiate translation. The *src* AUG initiation codon lies 3′ of the leader segment, 90 bases downstream of the *gag* initiation codon in the spliced *src* message. However, in the spliced *src* message a UGA termination codon lies between the *gag* AUG and the *src* AUG. All three codons are in the same reading frame. By using oligonucleotide-directed mutagenesis, the UGA termination codon has been converted to CGA. Cells infected with the mutant (called 1057 CGA) were spindle shaped, distinct from the rounded shape of cells infected with the parental Rous sarcoma virus. The mutant virus initiates *src* ranslation at the *gag* AUG, producing a 63,000-dalton *src* protein. We suggest that the wild-type *src* message produces two polypeptides, a very small (nine-amino acid) peptide that is initiated at the *gag* AUG and the 60,000-dalton *src* protein that is initiated at the *src* AUG.

Rous sarcoma virus (RSV) is apparently transcribed into a single RNA species beginning at the left end of the R sequence in the left-hand long terminal repeat and terminating at the right end of the R sequence in the right-hand long terminal repeat. This transcript constitutes genomic RNA and serves as the precursor for RSV messenger RNAs. Four messenger RNAs are known: *gag* mRNA, which is apparently unspliced and is thought to be identical to genomic RNA; *gag-pol* mRNA, which may have a small region near the *gag-pol* junction removed by RNA splicing; *env* mRNA, which has most of *gag* and *pol* removed by splicing; and *src* mRNA, which has most of *gag*, *pol*, and *env* removed by splicing (30; Fig. 1). The splice donor for the *env* and *src* mRNA lies five codons downstream of the AUG used to initiate the translation of *gag* (14, 26). All four mRNAs contain the same 5′ untranslated leader and the AUG used to initiate *gag*. The *gag* AUG is used to initiate the translation of *gag*, *gag-pol*, and *env* mRNAs (14, 15, 26). However, *src* mRNA is initiated at an AUG that lies 90 bases downstream from the *gag* AUG in the spliced *src* mRNA (8, 26, 28, 29). Why does not *src* mRNA initiate at the *gag* AUG?

In *src* mRNA there is an in-frame UGA codon that lies 24 bases downstream of the *gag* AUG. In the genome, this UGA codon lies nine bases downstream of the splice acceptor site between *env* and *src*; in the SR-A strain of RSV, the UGA codon is at position 1057 (numbering from the *Eco*RI site in *env*; 8, 26, 29; Fig. 2). If there is translational initiation at the *gag* AUG in the normal *src* message, the expected product would be a peptide nine amino acids long. Whether or not this peptide is synthesized, there is reasonably efficient initiation at the *src* AUG, 63 bases downstream of the UGA termination codon. In this report we have examined the position of initiation of *src* translation in a mutant virus that has the UGA codon at position 1057 converted to CGA by site-directed mutagenesis. Since all three codons, the *gag* AUG, the UGA, and the *src* AUG, are in frame, initiation of *src* translation at the *gag* AUG in the mutant would result in the synthesis of a *src* protein of ca. 63,000 daltons. Cells infected with the mutant virus were found to express a 63,000-dalton *src* protein, indicating that in the mutant most, if not all, of the *src* protein is initiated at the *gag* AUG. The mutant virus still causes morphological changes in cultured fibroblasts; however, cells infected by the 1057 CGA mutant are spindle shaped, which clearly distinguishes them from cells transformed by the parental virus.

The simplest explanation of these data is that the wild-type *src* message is polycistronic. The small (nine amino acids in length) open reading frame that begins at the *gag* AUG is probably translated in wild-type *src* mRNA; in addition, the *src* protein is translated from the same message. The implications of these observations are given below.

MATERIALS AND METHODS

Cells. C/O chicken embryo fibroblasts (SPAFAS, Inc.) were grown in Dulbecco modified Eagle medium containing 5% calf serum, 5% fetal calf serum, and 0.3% (wt/vol) tryptose phosphate containing penicillin and streptomycin.

Chemicals and enzymes. The Klenow fragment of DNA polymerase I and restriction enzymes were purchased from New England Biolabs or Bethesda Research Laboratories, alkaline phosphatase was purchased from Bethesda Research Laboratories, T4 polynucleotide kinase was purchased from P-L Biochemicals, Inc., agarose and S1 nuclease were purchased from Sigma Chemical Co., acrylamide was purchased from Bio-Rad Laboratories, and DNA linkers were purchased from Collaborative Research, Inc.

Preparation of cloned retrovirus DNA. The construction of the virus used in these experiments, 882/910-2795/2795, has been described elsewhere (27; S. H. Hughes and E. Kosik, Virology, in press). Briefly, the large circular form of integrated DNA from the SR-A strain of RSV (9) was cloned into the *Sal*I site of a pBR322 derivative that lacks the *Cla*I site. *Cla*I linkers were introduced either at restriction sites or at the ends of DNA fragments produced by digestion with *Bal*31 and repaired with the large fragment (Klenow fragment) of *Escherichia coli* DNA polymerase I. The precise location of the linkers was determined by DNA sequencing (21).

The *src*-containing segment of the virus, 910-2795, is

* Corresponding author.
† Present address: NCI-Frederick Cancer Research Facility, Frederick, MD 21701.

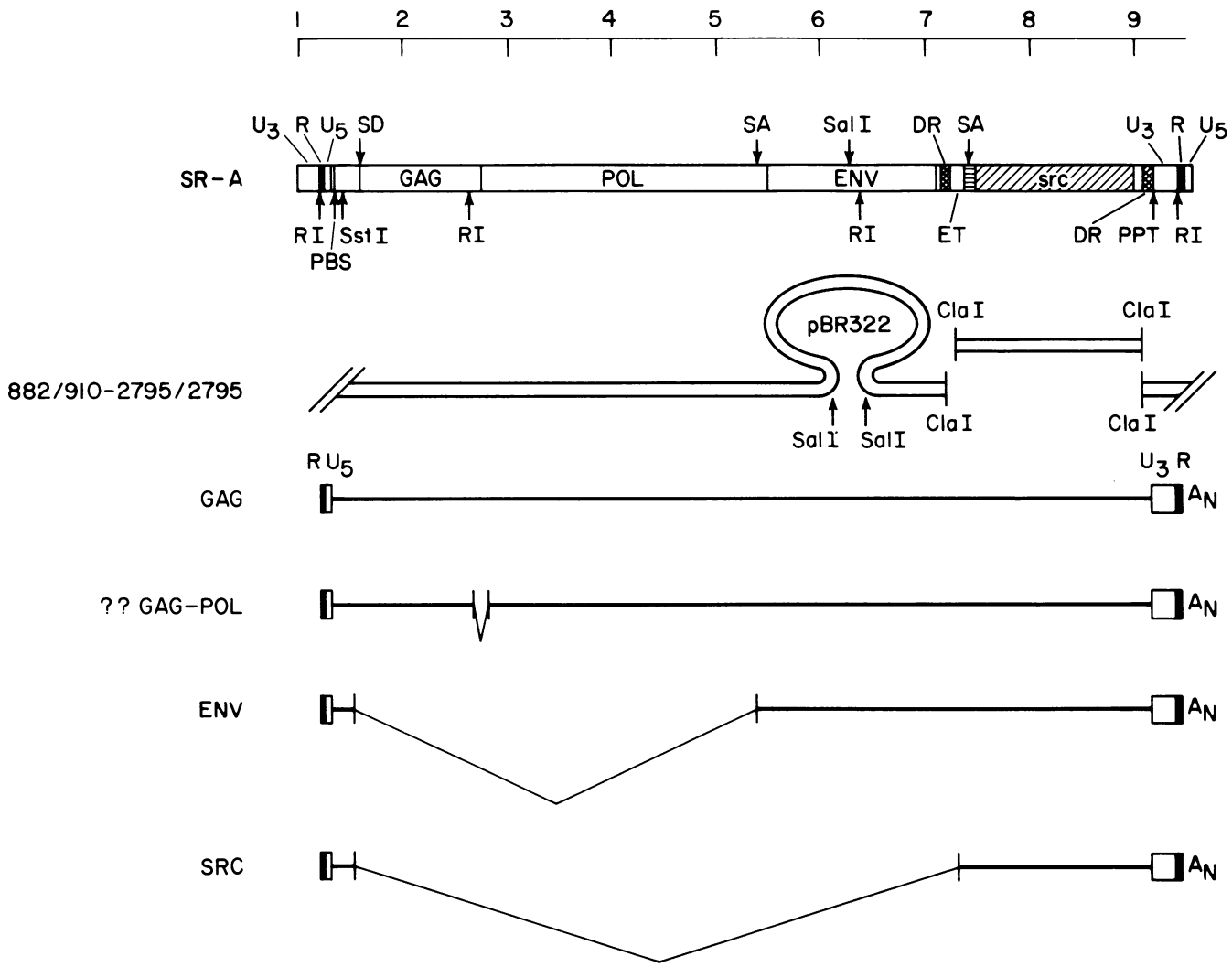

FIG. 1. RSV genome and its mRNAs. At the top of the figure is a scale in kilobases. Just below this scale is a diagram of the DNA of the wild-type SR-A strain of RSV. The genes, *gag*, *pol*, *env*, and *src*, are shown on the drawing, as are structural features like the splice donor (SD), splice acceptors (SA), U_3, R, U_5, primer binding site (PBS), direct repeats (DR), polypurine tract (PPT), and sequence of unknown origin (ET) that lies between the virally derived sequences and those derived from *c-src*. The sites for enzymes *Eco*RI (RI), *Sal*I, and *Sst*I are also marked. Below the diagram of the wild-type SR-A virus is a schematic drawing of the 882/910-2795/2795 plasmid. This construction is circular so the ends are undefined. In the plasmid, these ends are joined to each other. Plasmid pBR322 has been inserted in the *Sal*I site, and *Cla*I linkers have been inserted between *env* and *src* and between *src* and the long terminal repeat. The virus is named for the position of the *Cla*I linkers relative to the *Eco*RI site in *env*. Below the drawing of the 882/910-2795/2795 virus are drawings of the four RSV messages *gag*, *gag-pol*, *env*, and *src*. Regions removed by splicing are shown as thin lines between vertical bars. The *gag-pol* message is marked by ?? since the splice indicated in the drawing has been conjectured but not demonstrated.

flanked by *Cla*I sites and was subcloned into the *Cla*I site of pBR322. A clone was selected in which the upstream end of the *src* segment (position 910) was close to the *Eco*RI site in pBR322. The *Eco*RI-to-*Sma*I segment of this plasmid, which contains the *Eco*RI-to-*Cla*I segment of pBR322 and the *Cla*I-to-*Sma*I segment of the *src* subclone (910-1627), was cloned into the M13 vector Mp11. Appropriate clones were identified by dideoxy sequencing of Lac^- M13 clones.

Oligonucleotide-directed mutagenesis. The 13-base-long oligonucleotide CTGAGCCGACTCT was synthesized by the phosphate triester method (10) with a Biosearch SAM DNA synthesizer and purified on a 20% acrylamide–urea gel. This sequence was chosen after a computer search demonstrated that there were no close matches in the sequence of the M13 vector and only one close match in the *src* sequence. The specificity of priming that was predicted by the computer was tested in two ways. First, the oligonucleotide was used to prime DNA synthesis from the M13 clone that contained the 910-1627 *src* segment, and the resulting double-stranded DNA product was digested with various restriction enzymes. Analysis of these digestion products on an acrylamide-urea gel demonstrated that priming occurred only at the appropriate site in the clone (data not shown). In addition, the synthetic oligonucleotide was used as a primer in a dideoxy-sequencing reaction. This also showed that the oligonucleotide primed DNA synthesis only at the appropriate site on the M13 Mp11 clone that contained the 910-1627 *src* segment (data not shown). Mutagenesis was done by the protocols of Zoller and Smith (32). The mutagenic oligonucleotide was used as a primer in an in vitro reaction containing the Klenow fragment of DNA polymerase I and T4 DNA ligase. In this reaction a small portion of the single-

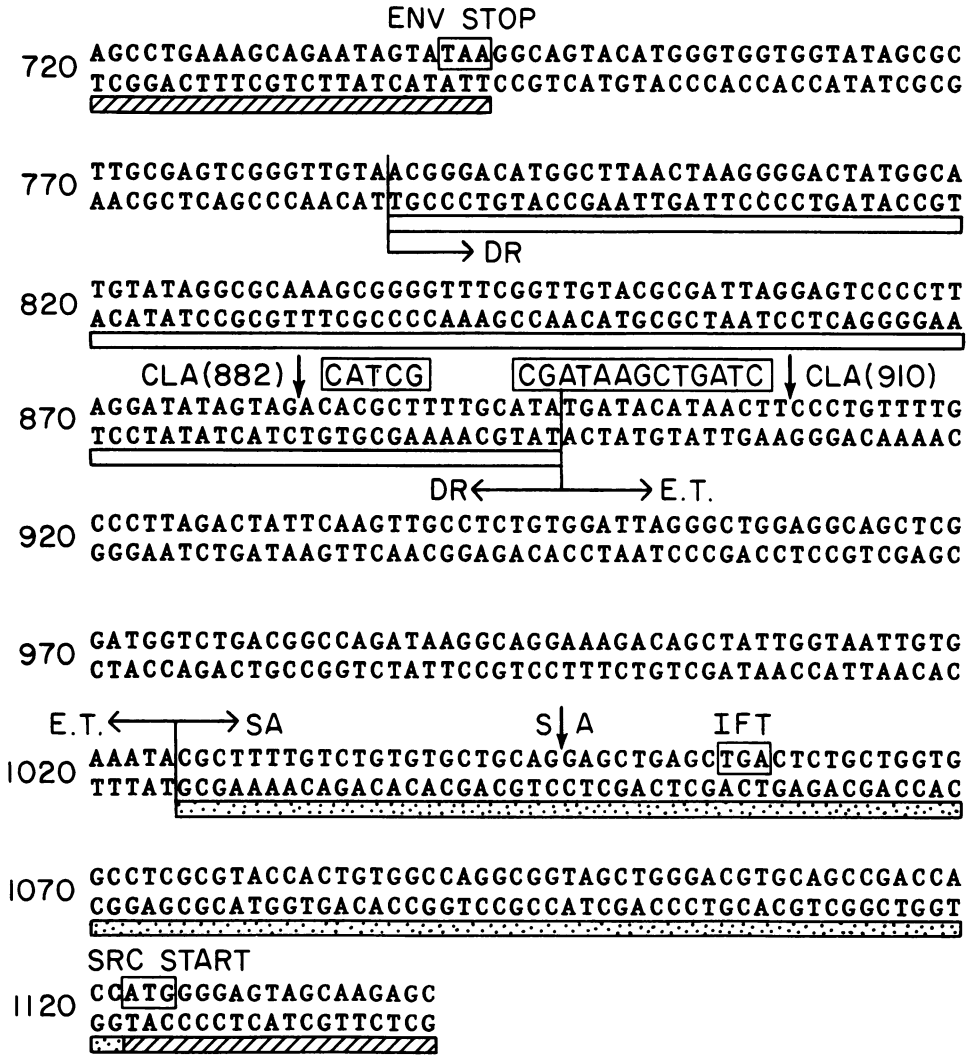

FIG. 2. Sequence of the region between *env* and *src*. The sequence of the region between *env* and *src* is given; the sequence is numbered from the *Eco*RI site in *env*. The joint at 882-910 in the 882/910-2795/2795 virus can be deduced by joining the half-*Cla*I sites shown in the boxes above the wild-type sequence at positions 882 and 910. This joint is well upstream of the *src* splice acceptor (SA), which is at position 1048. The TGA codon at position 1057 is boxed and labeled (IFT).

stranded M13 DNA was converted to covalently closed circular double-stranded molecules. These covalently closed circular DNA molecules were purified by sedimentation in alkaline sucrose gradients and used to transform *E. coli*.

Thirty-six clones were picked at random and grown as 3-ml cultures. A 500-μl volume of supernatant from each of these cultures was filtered through a nitrocellulose sheet, using a 96-well apparatus. Each clone was filtered through two wells, so that there was a duplicate set of DNA samples on the nitrocellulose filter. The filter was dried and then processed for hybridization as described by Grunstein and Hogness (13). The filter was hybridized to the 13-base-long mutagenic oligonucleotide which had been labeled at its 5′ end with polynucleotide kinase. After 1 h of hybridization, the filter was washed in 6× SSC (0.9 M NaCl plus 0.09 M sodium nitrate) at 25°C and autoradiographed. The filter was then washed at successively higher temperatures and audioradiographed after each wash. After a wash at 44°C, 6 of the clones retained the mutagenic oligonucleotide much better than did the other 30. Two of the clones that gave the strongest signal were replaqued, and three plaques were chosen from each of the two clones and sequenced. All six isolates tested had the desired change: the T at position 1057 was converted to a C. One of these clones was used as a template for DNA synthesis with the M13 sequencing primer. In this reaction most of the insert region was converted to double strand, and this portion was excised by digestion with *Eco*RI and *Sma*I. This *Eco*RI-to-*Sma*I fragment was recloned into an *Eco*RI-*Sma*I partial digest of the original pBR322 clone that carried the 910-1627 *src* segment. To confirm that the mutagenized segment of *src* had been acquired, the 910-1627 segment from the reconstructed pBR322 clone was recloned into Mp11 and sequenced. The *src* insert (910-2795) from the mutagenized version of the pBR322 clone was excised with *Cla*I and introduced at the *Cla*I site of the 822/2795 vector. The resulting construction is identical to 882/910-2795/2795 except that the T at position 1057 has been converted to a C.

For transfection into chicken cells, cloned viral DNA was isolated from an *E. coli* culture after amplification with chloramphenicol and then banded twice in $CsCl_2$-ethidium bromide gradients. The DNA was extracted several times

with isopropanol, once with a 1:1 phenol-chloroform mixture, diluted 10-fold, and ethanol precipitated.

Transfection into chicken cells. Before transfection, permuted viral DNA was released from plasmid DNA by digestion with the restriction endonuclease *Sal*I, the site into which the permuted viral DNA had been cloned, and ligated with T4 DNA ligase at a DNA concentration of 500 μg/ml. The ligation products were purified by phenol extraction and precipitated with ethanol. Approximately 5 μg of viral DNA was transfected into 30% confluent chicken embryo fibroblasts on a 10-cm plate by using modifications (31) of the technique of Graham and van der Eb (12). At 4 h after transfection, culture medium was removed, and the cells were exposed to 2 ml of culture medium containing 15% (vol/vol) glycerol for 2 min. The medium containing glycerol was removed, and fresh culture medium was added. Cells were split as necessary, and 10 to 12 days later, 10 ml of culture medium was assayed for reverse transcriptase activity (24) with a polyribocytidylate template, an oligodeoxyguanidylate primer, and [α-^{32}P]dGTP at a concentration of 10 μM and a specific activity of 1 Ci/mmol.

DNA sequencing. DNA samples were prepared from individual M13 clones and sequenced by the dideoxy-chain terminator method of Sanger and Coulson (25).

Antisera. Sera from tumor-bearing animals was obtained by injection of newborn animals with the Schmidt-Ruppin strain of RSV as described previously (3). Antiserum directed against pp60^{src} produced in *E. coli* (11) was obtained from R. L. Erikson, Harvard University, Boston, Mass.

Immunoprecipitation. Cells were radiolabeled as described in each figure legend and lysed in RIPA buffer as described previously (3). After clarification (49,000 × *g* for 30 min), the lysate was incubated with antibody for 30 min. The immune complexes were absorbed to Formalin-fixed *Staphylococcus aureus* (17) and washed three times with RIPA buffer. The immunoprecipitated proteins were eluted with electrophoresis sample buffer (3) and separated on 7.5% sodium dodecyl sulfate (SDS)-polyacrylamide gels. [^{35}S]methionine-labeled gels were subjected to fluorography as described previously (1) and exposed to X-Omat 5 film (Kodak).

Immune-complex protein kinase assay. Cell lysates were prepared from 5 × 10^6 cells and incubated with antiserum, and the resulting immune complexes were prepared as described above. After three detergent washes, the immunoprecipitates were washed with 0.15 M NaCl–10 mM Tris-hydrochloride (pH 7.2)–5 mM $MgCl_2$ and incubated with 30 μl of 10 mM Tris-hydrochloride (pH 7.2)–5 mM $MgCl_2$–5 μCi of [α-^{32}P]ATP (carrier free; ICN Pharmaceuticals Inc.) for 10 min at 0°C (7, 20). The reaction was terminated by the addition of electrophoresis sample buffer. The reaction products were analyzed on 7.5% SDS-polyacrylamide gels.

Partial proteolytic peptide mapping. Protein bands were excised from unfixed gels and subjected to electrophoresis on a 12.5% SDS-polyacrylamide gel in the presence of 50 ng of *S. aureus* V8 protease (Sigma Chemical Co.) as described by Cleveland et al. (4).

RESULTS

Mutagenesis of the RSV genome was performed on a subgenomic fragment that was cloned in an M13 vector. The *src* clone used in the construction was derived from a cloned copy of the SR-A strain of RSV originally constructed by DeLorbe and co-workers (9). Because the SR-A strain of RSV does not have convenient cloning sites near *src*, a derivative (882/910-2795/2795) (Fig. 1), which has *Cla*I sites in the noncoding regions flanking *src*, was used (Hughes and Kosik, in press). In this construction the *Cla*I linker that lies between *env* and *src* is upstream of the splice acceptor, and as a consequence, the linker is removed from the *src* message produced by this construction (cf. Fig. 2 and 3). Cells infected with 882/910-2795/2795 display the morphology typically associated with cells transformed by the SR-A strain of RSV and produce the expected 60,000-dalton *src* protein (see Fig. 5). The *src* segment of the construction (910-2795) was subcloned into the *Cla*I site of pBR322, with the insert oriented so that the upstream portion of *src* was adjacent to the *Eco*RI site of pBR322. The *Eco*RI-to-*Sma*I segment of this plasmid was subcloned into the M13 vector Mp11. The resulting M13 clone contains the antisense strand of a segment (910-1627) that includes the *src* splice acceptor, the UGA termination codon at 1057, and the *src* AUG. A 13-base-long mutagenic oligonucleotide was prepared to convert the UGA codon at position 1057 to CGA (arginine). By using the oligonucleotide as a primer, the single-stranded M13 DNA containing the 910-1627 region was converted to a double-stranded closed circular DNA molecule. This DNA was used to transform *E. coli*, and M13 clones were recovered that had the UGA at position 1057 converted to CGA (Fig. 4). A double-stranded copy of this region was prepared in vitro by using the M13 sequencing primer. The 910-1627 segment was reintroduced into an *Eco*RI-*Sma*I partial digest of the original pBR322 910-2795 clone. Clones containing the CGA codon at position 1057 were isolated, and the sequence of the reconstructed pBR322 clones were confirmed by recloning the *Eco*RI-to-*Sma*I segment in M13 Mp11 and sequencing. The 910-2795 *Cla*I fragment was reintroduced into the *Cla*I site of the 882/2795 retrovirus vector, reconstituting the 882/910-2795/2795 construction, with the conversion of UGA to CGA at position 1057.

DNA was prepared from this construction (called 1057 CGA), digested with *Sal*I to remove the plasmid vector, and concatemerized by ligation. The DNA was introduced into chicken embryo fibroblasts by calcium phosphate transfection. Approximately 10 days later foci appeared on the plate. Cells in the foci induced by the 1057 CGA mutant were spindle shaped, distinct from the cells in foci induced by the parental 882/910-2759/2795 virus, which are round, typical of cells transformed by the SR-A strain of RSV (Fig. 5). Cells infected and transformed by the 1057 CGA mutant were analyzed for viral *src* protein both by the immunoprecipitation of protein from cells metabolically labeled with ^{32}P (3, 5, 6; Fig. 6) and by the *src*-dependent phosphorylation of the heavy chain of immunoglobulin G (7, 20; Fig. 7). By both criteria, cells infected with the 1057 CGA mutant contained viral *src* protein. However, the bulk of the viral *src* protein present in the cells infected with the 1057 CGA mutant migrates on SDS-polyacrylamide gels at an apparent mass of 63,000 daltons, significantly larger than the protein specified by the wild-type virus (3; Fig. 6). If translation of the *src* gene in the 1057 CGA virus is initiated at the AUG donated by splicing from the *gag* gene, the predicted *src* protein would contain an additional 31 amino acids upstream of the usual amino terminus of the *src* gene. Analysis of the mutant protein by a partial digest with *S. aureus* V8 protease (4) demonstrated that the amino-terminal peptide (V1) from this protein migrates more slowly than the analagous peptide from the 60,000-dalton *src* protein (Fig. 6). When the wild-type *src* protein is cleaved by V8 protease, the initial cut divides the protein into an amino-terminal fragment (V1) and a carboxy-terminal fragment (V2). Additional protease treatment causes the amino-terminal fragment to be cut into two smaller fragments (V3 and V4; Fig. 6). When the 63,000-

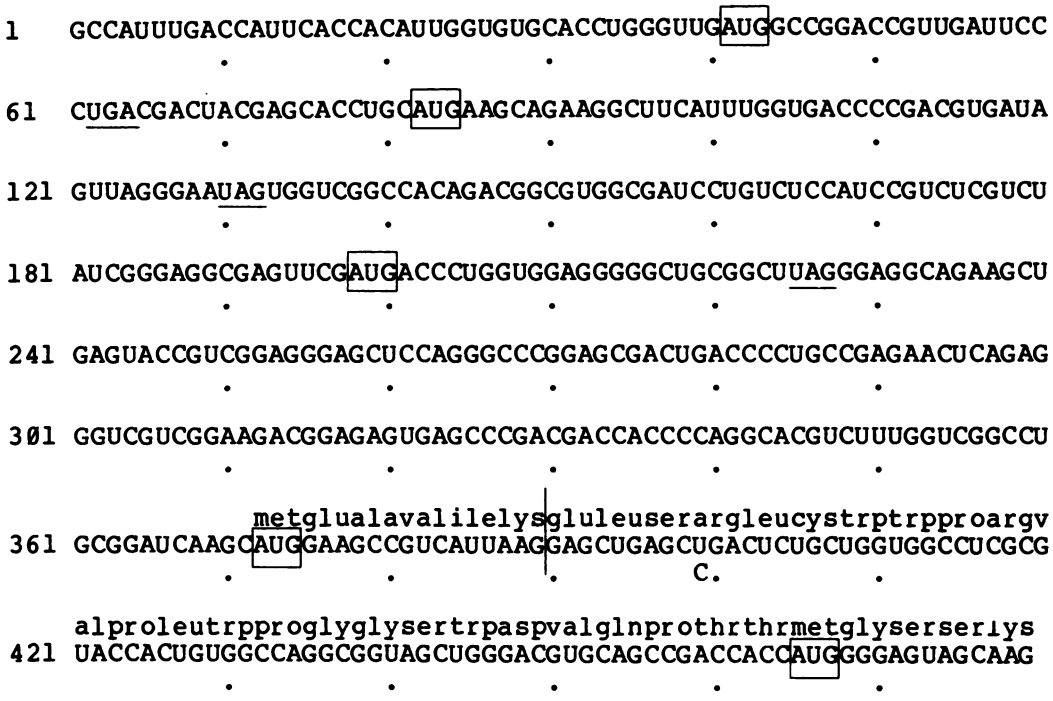

FIG. 3. The 5′ end of the *src* message. The sequence of the 5′ end of the *src* message is given, beginning with the cap nucleotide. The splice junction is at position 389 and is marked by a vertical line. The five AUGs in this position of the message are boxed. The first three AUGs are closely followed by termination codons; these termination codons are underlined. The 1057 CGA mutation results in a change from U to C in the message at position 399. The U, which is present in the wild-type message, is shown on the line as part of the RNA sequence; the C, which is present in the 1057 CGA mutant, is shown underneath the U. The 1057 CGA mutant initiates protein synthesis at the gag AUG (position 372); the inferred protein sequence of the amino terminus of the 63,000-dalton *src* protein made by the 1057 CGA mutant is shown above the RNA sequence.

dalton *src* protein from 1057 CGA is digested with V8 protease, the V1 fragment is larger than the V1 fragment from the wild-type *src* protein. The V1 fragment from the 63,000-dalton protein is not readily cleaved to fragments V3 and V4, presumably because of the altered structure of the V1 fragment. This indicates that the amino-terminal end of the *src* protein has been altered, providing strong evidence that the *src* protein is initiated at the AUG donated from the *gag* gene.

Cells infected with 1057 CGA contained, in addition to the larger mutant protein of 63,000 daltons, various amounts of a 60,000-dalton *src* protein that is either very similar or identical to the wild-type *src* protein in its migration on SDS-polyacrylamide gels. The possibility that this protein is the cellular homolog of the *src* protein was ruled out since this 60,000-dalton protein was precipitated by an antiserum that does not recognize the *c-src* protein. The 60,000-dalton protein could be a cleavage product of the 63,000-dalton protein, could be the product of initiation at the *src* AUG in the mutant 1057 CGA, or could result from reversion of the 1057 CGA mutation. To test the possibility that the 60,000-dalton *src* gene product was the product of the processing of the 63,000-dalton protein, cells were pulse-labeled with [^{35}S]methionine, the label was removed, and the cells were allowed to grow for various periods in the presence of nonradioactive methionine. The *src* proteins were immunoprecipitated and separated by SDS-polyacrylamide gel electrophoresis (Fig. 8). In this experiment only a small amount of the 60,000-dalton *src* protein was synthesized during the pulse, and the ratio of the 63,000- to the 60,000-dalton protein did not seem to change significantly during the chase. This suggests that the 60,000-dalton protein is not a

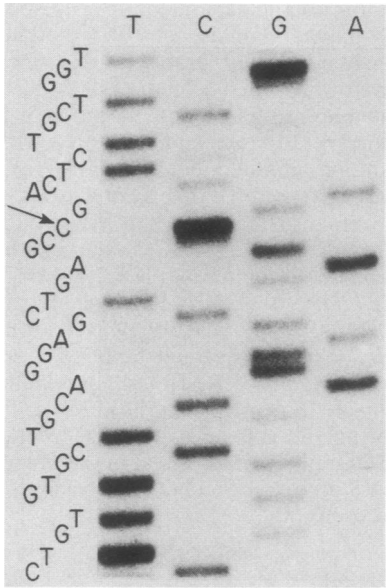

FIG. 4. Sequence of the 1057 CGA mutant. Dideoxy sequencing was done with the M13 Mp11 clones that were isolated after oligonucleotide-directed mutagenesis. Sequencing reactions were run by the protocols of Sanger and Coulson (25), and the samples were analyzed on a 5% acrylamide–urea gel. The figure shows the portion of the autoradiograph around position 1057. In this clone the base at position 1057 (marked with an arrow) has been converted to a C.

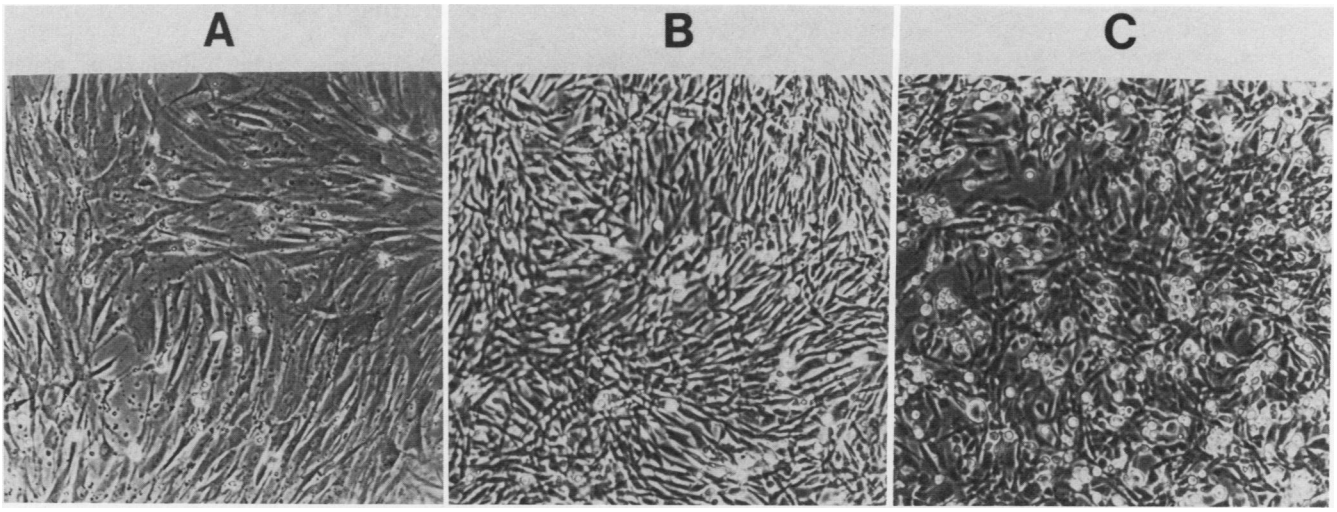

FIG. 5. Morphology of normal chicken cells and cells infected with the parental virus. Chicken cells were transfected with DNA from the 1057 CGA mutant or the parental virus (882/910-2795/2795) (see the text). When reverse transcriptase levels peaked, about 3 weeks after transfection, photomicrographs were taken of represenative areas of the infected cultures and the uninfected controls. (A) Cells from a control culture; (B) cells from a culture infected with the 1057 CGA mutant; and (C) cells from the parental virus 882/910-2795/2795.

proteolytic cleavage product derived from the 63,000-dalton protein. We cannot distinguish among the remaining possibilities. There is, however, a tendency for the 60,000-dalton protein to be more prominent in virus stocks that have been passaged several times, suggesting that at least some of the 60,000-dalton protein may be derived from revertant virus.

Since the effects of mRNA secondary structure on the initiation of translation may be subtle and are not well understood, it is impossible to be certain that the effect of 1057 CGA mutation is not in part or wholly the result of an effect on the secondary structure of the *src* mRNA. However, the observation that two other mutations that would be expected to have a much more dramatic effect on the structure of the *src* message do not alter the site at which *src* translation initiates suggests that this is an unlikely possibility.

There are two viruses that have substantial modifications in the region between the UGA codon at position 1057 and the *src* AUG at position 1123 (Hughes and Kosik, in press). The first mutant, 779/989-1089/1079-2795/2795, has a 10-base duplication just downstream of the UGA codon at position 1057. The duplicated regions are joined by a synthetic DNA segment 12 bases in length. Despite this 22-base insertion, *src* translation is initiated at the *src* AUG. It should be noted that the insertion of 22 bases changes the relationship between the UGA codon at 1057 and the *src* AUG at 1123. These two codons, which are in the same reading frame in the wild-type virus, are in different reading frames in the 779/989-1089/1079-2795/2795 mutant. The second mutant, 882/1089-989/1079-2795/2795, has the segment that contains the *src* splice acceptor (989-1089) inserted in the antisense orientation. This inverted segment, when linked to the 1079-2795 segment, creates a new splice acceptor site at position 1079. The *src* message produced by this virus is missing the segment from 1048 to 1079. This missing segment includes the 1057 UGA codon. However, the splicing event in the 882/1089-989/1079-2795/2795 virus brings the UGA codon in frame with the *gag* AUG in the *src* message (Hughes and Kosik, in press). This virus also specifies the 60,000-dalton form of *src*. In the spliced *src* message produced by this virus the *gag* and the *src* AUGs are in different reading frames.

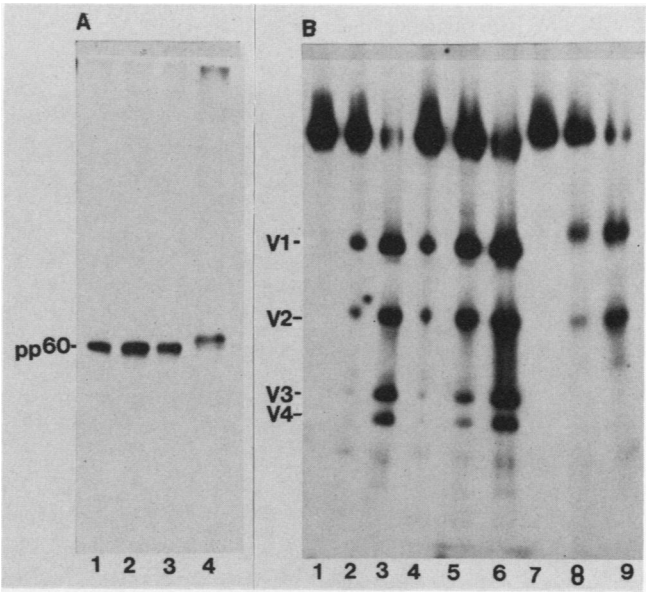

FIG. 6. Analysis of the *src* protein expressed in cells infected with mutant variants of RSV. (A) Cell cultures were labeled for 4 h with ^{32}P, and detergent lysates were prepared (3). *src* was immunoprecipitated from cells lysates by using 1 μl of antiserum directed against *src* expressed in *E. coli* (11) as described in the text. The viruses used for infection of the cultures were 882/1089-989/1079-2795/2795 (lane 1), 779/989-1089/1079-2795/2795 (lane 2), 882/910-2759/2795, the parental virus (lane 3), and 1057 CGA (lane 4). (B) The *src* protein from cells infected with the wild-type SR-A virus and the 63,000-dalton protein from 1057 CGA-infected cells were precipitated with anti-*src* serum, excised from a 7.5% gel, and subjected to electrophoresis in the presence of V8 protease on a 12.5% gel as described previously (4). Lanes 1, 2 and 3 contain *src* from cells infected with parental virus (882/910-2795/2795); lanes 4, 5, and 6 contain the 60,000-dalton protein from cells infected with 779/989-1089/1079-2795/2795; and lanes 7, 8, and 9 contain the 63,000-dalton protein from cells infected with 1057 CGA. Lanes 1, 4, and 7, no enzyme; lanes 2, 5, and 8, 20 ng of enzyme; lanes 3, 6, and 9, 50 ng of enzyme.

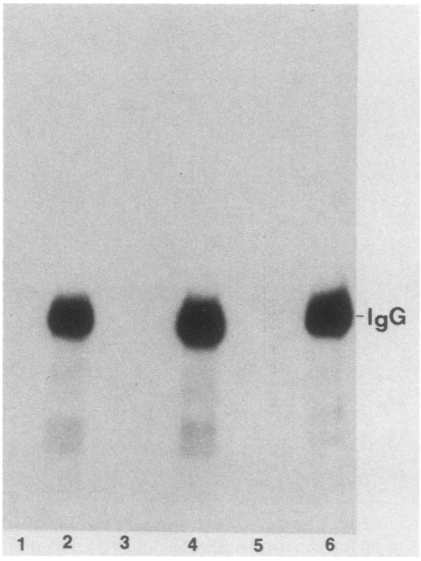

FIG. 7. Immune complex protein kinase activity of the *src* protein from cells infected with RSV mutants. Immunoglobulin phosphorylation was assayed as described in the text after precipitation of *src* from cells infected with the parental virus (882/910-2795/2795) (lanes 1 and 2) and with the mutants 779/989-1089/1079-2795/2795 (lanes 3 and 4) and 1057 CGA (lanes 5 and 6), using normal rabbit serum (lanes 1, 3, and 5) or tumor-bearing animal serum (lanes 2, 4, and 6). IgG, Immunoglobulin G.

DISCUSSION

The data presented here indicate that a single base change 63 bases upstream of the *src* AUG virtually eliminates initiation at the *src* AUG; in the 1057 CGA mutant, essentially all initiation occurs upstream at the *gag* AUG. Other mutants, one with a 22-base insertion in the *src* message, and the other with a 31-base deletion, initiate *src* translation at the *src* AUG. It is therefore unlikely that the effect seen in the 1057 CGA mutant is a result of the alteration in the secondary structure of the *src* mRNA. To explain the data, we must account for an apparent shift of 90 bases in the site of initiation of translation in two *src* messages that differ by a single base. It is unclear how the change could activate the *gag* AUG if that codon were not normally used as a site of initiation. We therefore suggest that translation of the wild-type *src* message produces a nine-amino acid polypeptide in addition to the 60,000-dalton *src* protein. There are several examples of single mRNAs that appear to have more than one functional site of initiation (2, 16, 21, 22), and two recent reports support the idea that termination codons can have a profound influence on initiation (19a, 20a).

Although the scanning hypothesis (18, 19) adequately explains the initiation of protein synthesis for most eucaryotic messages, it fails to explain initiation on the *src* message. Any proposed explanation for *src* initiation must be able to account for the observation that in the 1057 CGA mutant virus the *src* AUG is used only poorly, if at all.

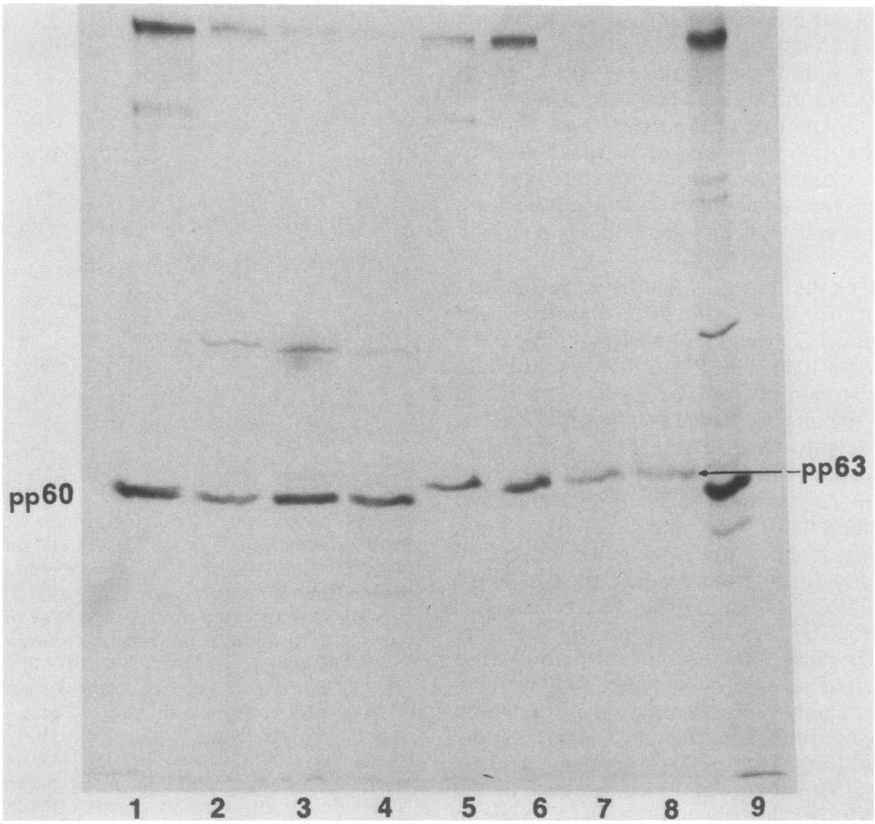

FIG. 8. Analysis of the fate of the *src* protein. Cells transformed by the parental virus (882/910-2795/2795) (lanes 1 through 4) or the 1057 CGA mutant virus (lanes 5 through 9) were labeled for 5 min with 275 μCi of [^{35}S]methionine per ml. The cells were then washed with media containing unlabeled methionine and either harvested for immunoprecipitation (lanes 1 and 5) or incubated for 30 min (lanes 2 and 6), 2 h (lanes 3 and 7), or 4 h (lanes 4 and 8) in unlabeled medium. The *src* protein was immunoprecipitated from lysates with antiserum prepared against *src* expressed in *E. coli*.

Two simple models can be proposed; however, neither model accounts for all of the available data. First, it is possible that all of the ribosomes load onto the *src* message at or near the 5′ end and scan downstream for the first acceptable AUG (18, 19). Alternatively, the ribosomes could, for certain messages like *src*, load independently at more than one AUG. It should be noted that there are three AUG codons upstream of the *src* AUG in *src* mRNA, although the *gag* AUG is the first AUG to conform to the rules proposed by Kozak (18, 19). All three of the AUGs that lie upstream of the *gag* AUG are followed closely by an in-frame terminator. If it is proposed that the ribosomes all load at or near the 5′ end of the *src* message, then in the case of the wild-type *src* message, initiation would take place at the *gag* AUG and termination would take place at the UGA codon at 1057, but the ribosomes would stay on or near the *src* mRNA and would reinitiate at the *src* AUG. If this model is correct and the ribosomes stay on the message, the analysis of the two mutants (779/989-1089/1079-2795/2795 and 882/1089-989/1079-2795/2795) that have frame shifts between the termination codon and the *src* AUG demonstrates that the ribosomes do not keep track of the translational reading frame in the region between the termination codon and the second AUG.

However, experiments attempting to determine at which AUG robosomes initially bind when virion RNA is translated in vitro suggest that the very first AUG is the preferred site, despite the fact that this AUG does not have the adjacent bases common to most initiation codons (R. Peterson, C. Hensel, and P. Hackett, personal communication). If the experiments examining the initiation of the translation of virion RNA in vitro apply to the initiation of the translation of the *src* message in vivo, then there may actually be four small cistrons upstream of *src* and the ribosomes would load at or near the 5′ end of the message.

The alternatives to permitting ribosomes to load only at or near the 5′ end of the message are models that allow ribosomes to load onto the *src* mRNA independently at more than one AUG codon. Such models must be able to account for the failure to efficiently initiate at the *src* AUG in the 1057 CGA mutant virus. In this model it can be proposed that, in the mutant virus, ribosomes initiating at the *gag* AUG would, as they move down the message, block the access of ribosomes that would otherwise be able to initiate at the *src* AUG. The *src* AUG would be accessible in the wild-type message because, with wild-type *src* mRNA, the ribosomes that initiated at the *gag* AUG would fall off the message at the terminator and would not interfere with initiation at the *src* AUG.

The analysis of the in vivo translation products of certain *src* mutants suggests that in some cases ribosomes can choose to initiate at two alternative AUGs not separated by a termination codon (22). In these cases the first AUG can apparently be bypassed part of the time. These separate initiations on *src* mRNA in vitro seem to favor some sort of independent entry of ribosomes onto the message. The experiments looking at the initiation of translation of RSV virion RNA in vitro seem, however, to favor a scanning model.

The data presented here, taken together with the other available information, suggest that there can be multiple initiation sites on a single eucaryotic mRNA and that there may be a true polycistronic message in higher eucaryotes. In addition, the data presented here demonstrate that termination codons in the 5′ untranslated region of mRNAs can have profound effects on translation of the mRNAs. None of the current models for the initiation of translation in eucaryotes seem to adequately describe all of the data, and these theories may require modification.

The 1057 CGA mutant is also of interest because it produces an abnormal *src* protein. Although the 63,000-dalton *src* protein is capable of phosphorylating immunoglobulin G in the immune-complex kinase assay, cells infected with the 1057 CGA mutant do not have the morphological characteristics of cells transforming by the wild-type SR-A strain of RSV. We are attempting to determine how the additional amino acids affect the interactions of the protein with the host cell.

ACKNOWLEDGMENTS

This research was supported by Public Health Service grants CA 27951 and CA 2956900 from the National Institutes of Health and by grant MV 187 from the American Cancer Society.

LITERATURE CITED

1. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium-labeled protein and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
2. **Bos, J. L., L. J. Polder, R. Bernardo, P. Schrier, P. J. Van den Elsen, A. S. Van der Eb, and H. van Ormondt.** 1981. The 2.2 kd E1b mRNA of human Ad 12 and 5 codes for two tumor antigens starting at different AUG triplets. Cell **27**:121–131.
3. **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.
4. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecylsulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1102–1106.
5. **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.
6. **Collett, M. S., E. Erikson, and R. L. Erikson.** 1979. Structural analysis of the avian sarcoma virus transforming protein: sites of phosphorylation. J. Virol. **29**:770–781.
7. **Collett, 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.
8. **Czernilofsky, A. P., A. D. Levinson, H. E. Varmus, J. M. Bishop, E. Tisher, and H. M. Goodman.** 1983. Corrections to the nucleotide sequence of the *src* gene of Rous sarcoma virus. Nature (London) **301**:736–738.
9. **DeLorbe, W. J., P. A. Luciw, H. M. Goodman, H. E. Varmus, and J. M. Bishop.** 1980. Molecular cloning and characterization of avian sarcoma virus circular DNA molecules. J. Virol. **36**:50–61.
10. **Efimov, V. A., S. V. Reverdatto, and O. G. Chakhmakcheva.** 1982. New effective method for synthesis of oligonucleotides via phophotriester intermediates. Nucleic Acids Res. **10**:6675–6694.
11. **Gilmer, T. M., and R. L. Erikson.** 1983. Development of anti-pp60^{src} serum with antigen produced in *Escherichia coli*. J. Virol. **45**:462–465.
12. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
13. **Grunstein, M., and D. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. U.S.A. **72**:3961–3965.
14. **Hackett, P. B., R. Swanstrom, H. E. Varmus, and J. M. Bishop.** 1982. The leader sequence of the subgenomic mRNA's of Rous sarcoma virus is approximately 390 nucleotides. J. Virol. **41**:527–534.
15. **Hunter, E., E. Hill, M. Hardwick, A. Bhown, D. E. Schwartz, and R. Tizard.** 1983. Complete sequence of the Rous sarcoma

virus *env* gene: identification of structural and functional regions of its product. J. Virol. **46**:920–936.

16. **Jay, G., S. Nomura, C. W. Anderson, and G. Khoury.** 1981. Identification of the SV40 agnogene product: a DNA binding protein. Nature (London) **291**:346–349.
17. **Kessler, S. W.** 1975. Rapid isolation of antigens from cells with a staphylococcal protein A-antibody absorbent: parameters of the interaction of antibody-antigens complexes with protein A. J. Immunol. **115**:1617–1624.
18. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
19. **Kozak, M.** 1982. Analysis of ribosome binding sites from the S1 message of reovirus. J. Mol. Biol. **156**:807–820.

19a. **Kozak, M.** 1984. Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. Nucleic Acids Res. **12**:3873–3893.

20. **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.

20a. **Liu, C.-C., C. C. Simonsen, and A. D. Levinson.** 1984. Initiation of translation at internal AUG codons in mammalian cells. Nature (London) **309**:82–85.

21. **Lomedico, P. T., and S. J. Andrew.** 1982. Eukaryotic ribosomes can recognize preproinsulin initiation codons irrespective of their position relative to the 5′ end of mRNA. Nature (London) **299**:221–226.
22. **Mardon, G., and H. E. Varmus.** 1983. Frameshift and intragenic suppressor mutations in a Rous sarcoma provirus suggest *src* encodes two proteins. Cell **32**:871–879.
23. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
24. **Robinson, H. L.** 1976. Intracellular restriction on the growth of induced subgroup E avian type C viruses in chicken cells. J. Virol. **18**:856–866.
25. **Sanger, F., and A. R. Coulson.** 1975. A rapid method for determining sequencing in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94**:441–448.
26. **Schwartz, D., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous Sarcoma Virus. Cell **32**:853–869.
27. **Sorge, J., and S. Hughes.** 1982. Splicing of intervening sequences introduced into an infectious retroviral vector. J. Mol. Appl. Genet. **1**:547–559.
28. **Swanstrom, R., R. C. Parker, H. E. Varmus, and J. M. Bishop.** 1983. Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **80**:2519–2523.
29. **Takeya, T., R. A. Feldman, and H. Hanafusa.** 1982. DNA sequence of the viral and cellular *src* gene of chickens. I. Complete nucleotide sequence of an *Eco*RI fragment of recovered avian sarcoma virus which codes for gp 37 and pp60^{src}. J. Virol. **44**:1–11.
30. **Weiss, R. A., N. Teich, H. Varmus, and J. Coffin.** 1982. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **76**:1373–1376.
32. **Zoller, M., and M. Smith.** 1983. Oligonucleotide directed mutagenesis of DNA fragments cloned in M13 vectors. Methods Enzymol. **100**:468–500.