MOLECULAR AND CELLULAR BIOLOGY, Oct. 1983, p. 1746–1758
0270-7306/83/101746-13$02.00/0

Vol. 3, No. 10

Rous Sarcoma Virus Contains Sequences Which Permit Expression of the *gag* Gene in *Escherichia coli*

BRION MERMER, MICHAEL MALAMY, AND JOHN M. COFFIN*

Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111

Received 29 April 1983/Accepted 27 July 1983

Several aspects of Rous sarcoma virus gene expression, including transcription, translation, and protein processing, can occur within *Escherichia coli* containing cloned viral DNA. The viral long terminal repeat contains a bacterial promoter, and viral sequences at or near the authentic viral initiation codon permit the initiation of translation. These signals can direct the synthesis in *E. coli* of the viral *gag* gene precursor Pr76 or, when fused to a portion of the *lacZ* gene, a *gag*–β-galactosidase fusion protein. Pr76 is processed into *gag* structural proteins in *E. coli* in a process which is dependent upon the *gag* product p15. These observations suggest that *E. coli* can be used for the introduction and analysis of mutations in sequences relevant to viral gene expression.

The analysis of eucaryotic genes has been facilitated by their introduction into procaryotic cells. Many of the basic tools of bacterial genetics, such as those permitting selection and mapping of mutations, have not yet proven useful since most eucaryotic genes confer no phenotype on bacterial cells. This obstacle could be overcome to a large extent by fusing the eucaryotic gene of interest to a procaryotic gene which does have a readily detectable phenotype, such that expression of the procaryotic gene in bacteria is dependent upon expression of at least a portion of the silent gene. Such a procedure would be particularly suitable for the manipulation of genetic information which is essential in the replication cycle of animal viruses. We adapted such a system to the study of avian retroviruses by using fusion vectors which require that the inserted DNA provide active signals for the initiation of RNA and protein synthesis for the expression of bacterial β-galactosidase in *Escherichia coli*. In the course of this study, we have found, to our surprise, that a DNA copy of the Rous sarcoma virus (RSV) genome contains signals usable by *E. coli* for initiation of viral RNA and protein synthesis. This results in the synthesis of a precursor protein and its processing to yield a set of proteins much like those found in virus-infected avian cells.

Retroviruses replicate via a proviral DNA intermediate which is the product of reverse transcription of genomic RNA (30). The genetic structure of an avian sarcoma virus genome RNA may be denoted as 5′ R U_5 *gag pol env src* U_3 R 3′, and the DNA provirus may be denoted U_3 R U_5 *gag pol env src* U_3 R U_5, in which duplication of the U_3RU_5 regions during viral DNA synthesis creates a long terminal repeat (LTR) (15, 26). R denotes sequences which are reiterated at both ends of the RNA genome, whereas U_5 indicates noncoding sequences unique to the 5′ end. The U_3 region, which codes for no known product, contains essential sequences which probably include the promoter for viral RNA synthesis (34). The *gag* gene codes for a precursor protein of M_r 76,000 (Pr76), whose mRNA is thought to have the same structure as genome RNA (23). Pr76 is processed into the virion core structural proteins p19, p27, p12, and p15 (11). *pol*, *env*, and *src* encode reverse transcriptase, the virion envelope glycoproteins, and the transforming protein, respectively.

These genetic regions, with the exception of *src*, are essential for viral replication. U_3, by virtue of its role in determining the rate of viral RNA synthesis, may control viral growth rate (33), transformation by acquired *onc* genes (such as *src*) (3, 4), and leukemogenesis in the case of nontransforming viruses (22, 24, 32). The *gag* proteins, besides their structural function, may be required for other steps in viral replication. p15 has been implicated as a protease involved in the processing of Pr76 (36, 37). p19 binds preferentially to specific regions of viral RNA and also has an affinity for the lipid envelope (17). p12 is a basic protein which also binds to viral RNA, but without sequence specificity (11). The precise roles of individual *gag* proteins in replication have been difficult to determine genetically, since a single mutation may affect

not only that *gag* peptide but the extent to which other peptides are cleaved from Pr76 (35). In addition, the isolation of *gag* mutants is difficult since *gag* is required for viral replication (35).

Our approach toward manipulating the viral genome has been to take advantage of a previous report that the LTR of cloned RSV DNA contains a promoter signal capable of inducing the expression of the *tet* gene of the plasmid pBR322 in *E. coli* (20). The analysis and manipulation of procaryotic promoters have been simplified by the fusion of promoter sequences from the operon of interest to the structural gene for β-galactosidase (1). Plasmids which permit the formation of such fusions in vitro can also be used for the study of amino (N)-terminal coding sequences and translation regulatory signals within the inserted DNA, since the N terminus of β-galactosidase can be replaced with any of a large number of different protein sequences with retention of enzymatic activity (7). Simple culture techniques are available for the in vivo selection of regulatory and structural gene mutations which affect the synthesis of the fusion proteins. β-Galactosidase activity can also be readily quantitated in vitro, thus allowing the detection of mutations which affect the level of expression (19).

We found that several features of this genetic fusion system are adaptable to the study of avian retroviral genomes. First, sequences at or near those which signal the initiation of viral transcription and translation in infected avian cells can drive transcription and translation of the *lac* operon in *E. coli*. Second, when plasmids containing these sequences upstream of a complete *gag* gene were introduced into *E. coli*, we found that the *gag* gene product was synthesized and processed in a similar manner to that in avian cells. We used this phenomenon to provide genetic evidence that p15 plays a role in processing of the *gag* precursor. These findings indicate that several steps in the viral replication cycle, including transcription, translation, and processing, can be mimicked in bacteria and do not require systems unique to eucaryotic cells.

MATERIALS AND METHODS

Bacterial strains and plasmids. Plasmids were propagated and screened for β-galactosidase activity, and transcripts were analyzed in the *E. coli* K-12 derivative RV200 (*thi ΔlacX74 rpsL200*). Protein products were analyzed in the *deg* mutant BU8049 from A. Bukhari [*deg-2 Δ(lac)XIII rpsL*] (6). Plasmids pMC874 and pMC1403 (7) were from M. Casadaban. Plasmid pATV-8 was constructed by R. A. Katz and R. V. Guntaka (16). Recombinants between pATV-8 and pMC874 were constructed by standard techniques. The initial ligation of *Bam*HI fragments from pATV-8 with *Bam*HI-digested pMC874 produced many Lac$^+$ colonies. Analysis of plasmid DNA from Lac$^+$ colonies indicated that all contained multiple *Bam*HI fragments inserted into pMC874 with the 5.4-kilobase (kb) *Bam*HI fragment from pATV-8 always fused to *lac'Z* (Fig. 1). pBM1 was constructed from one such recombinant plasmid containing a tandem insertion of 5.4-kb fragments in direct repeat by digesting with *Bgl*II, which cuts once in the 5.4-kb fragment, and then religating.

Viruses and cells. Stocks of the Prague C strain of avian sarcoma virus (PrC RSV) were obtained by transfection of chicken embryo fibroblasts with cloned viral DNA derived from plasmid pATV-8; its identity was confirmed by T_1 oligonucleotide fingerprinting (P. Norton, personal communication).

For the preparation of labeled intracellular and virion proteins, PrC RSV was grown in turkey embryo fibroblasts as previously described (10).

The related viruses, nontransforming recombinant virus (NTRE-7), Rous-associated virus (RAV)-60, and RAV-2, have been previously described (9).

Purified avian myeloblastosis virus (AMV) virions were provided by R. Eisenman.

Detection of Lac$^+$ colonies. Bacterial colonies expressing the *lacZ* gene were detected on plates containing the indicator dye XG (5-chloro-4-bromo-3-indolyl-β-D-galactoside) or on lactose-MacConkey plates or on *lac* minimal plates at 37°C. For detection of *lacY* activity, melibiose minimal plates were used at 42°C (19).

DNA sequencing. DNA sequences of the RSV-*lacZ* fusion region of pBM3 were determined by the method of Maxam and Gilbert (18).

Preparation of cellular RNA and plasmid DNA. Total bacterial RNA was isolated after lysis of cells by freezing, thawing, and phenol extraction, as previously published (5), and treated with RNase-free DNase (Sigma Chemical Co.). The purified RNA was largely intact as determined by visualization of glyoxalated 16S and 23S rRNA with ethidium bromide after agarose gel electrophoresis (31). Total RNA from RAV-60- and NTRE-7-infected quail cells was a gift of S. Herman. 70S RNA from RAV-2 virions was prepared by K. Conklin.

Plasmid DNA was isolated in cesium chloride-ethidium bromide gradients.

S1 nuclease mapping. S1 mapping was performed as described previously (2), except that DNA probes were 5′ end labeled with polynucleotide kinase (38). Restriction fragments were purified by electrophoresis on neutral polyacrylamide gels and extracted from the gel by the crush-and-soak procedure (18), and the purification was monitored by the addition of a trace amount of end-labeled restriction digest. The eluted fragment was then labeled with ^{32}P at its 5′ ends (18). The optimum temperature for hybridization was determined by analysis of samples on neutral polyacrylamide gels after denaturation at 90°C for 10 min and incubation in hybridization buffer (80% deionized formamide–0.4 M NaCl–30 mM PIPES [piperazine-*N*,*N*′-bis(2-ethane sulfonic acid); pH 6.4]–1 mM EDTA) at various temperatures for 3 h. For localization of RNA 5′ ends, 8 to 10 μg of total cellular RNA was hybridized with 20,000 cpm of denatured probe at 50°C in 10 μl of hybridization buffer for 3 h. S1 digestions were carried out for 10 min at 42°C, with 10 U of S1. Carrier tRNA was added to control samples to maintain a consistent amount of total RNA. After digestion,

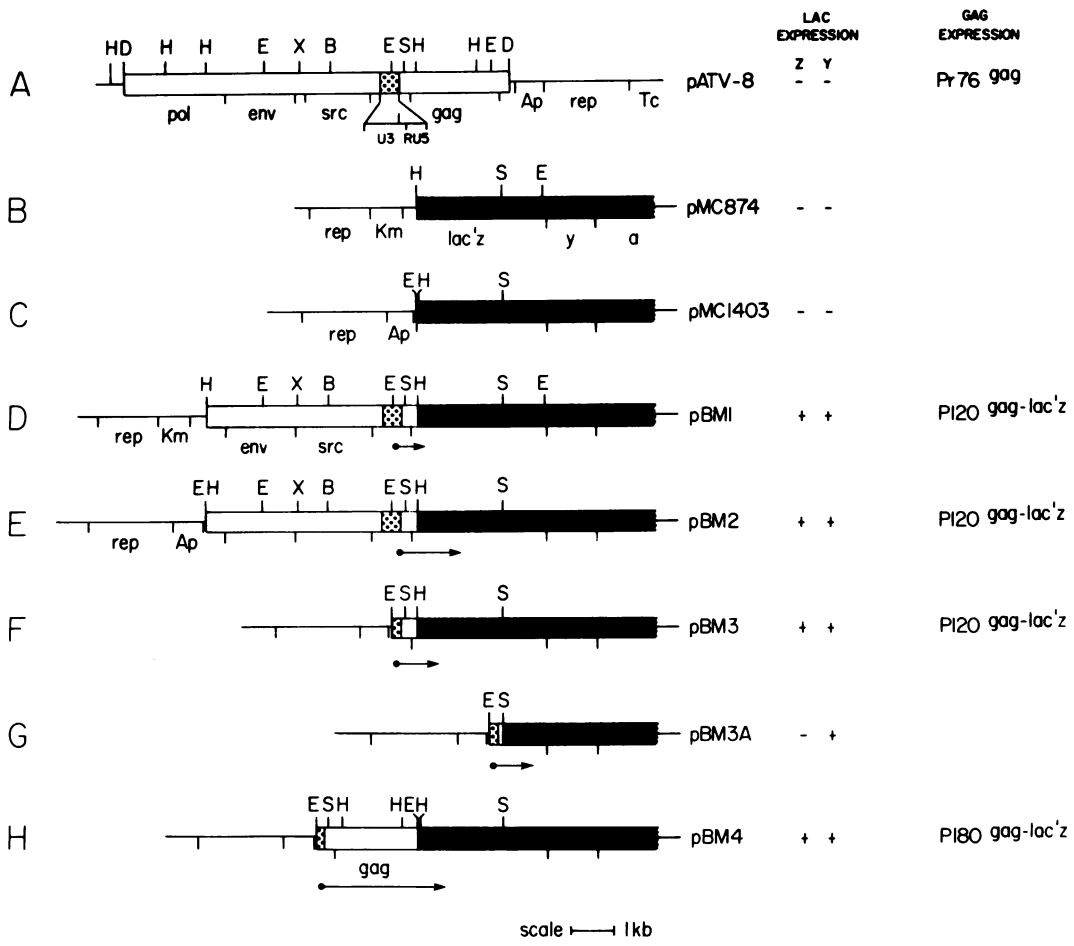

FIG. 1. Plasmids used in this study. All plasmids are drawn open at their unique *Sal*I sites. Relevant restriction sites are indicated above each plasmid (not all sites are shown): H, *Bam*HI; E, *Eco*RI; X, *Xho*I; B, *Bgl*II; S, *Sac*I; and D, *Hin*dIII. Approximate boundaries of various genetic markers are indicated below each plasmid. Expression of the *lacZ* and *-Y* cistrons is indicated by + or − to the right of each plasmid. Protein products initiating in *E. coli* at or near the *gag* initiation codon are listed under *gag* expression. The open bar indicates viral DNA, the cross-hatched bar indicates *lac* operon structural gene sequences, the line represents vector sequences, and the LTR is denoted by a dotted bar. (A) pATV-8. Unintegrated circular DNA of the PrC strain of RSV permuted at a *Hin*dIII site in *pol* and cloned at the *Hin*dIII site of pBR322. This clone contains only one LTR which is shown expanded threefold below the plasmid. (B) pMC874. *lac* fusion vector encoding resistance to kanamycin. *lac'Z* indicates the structural gene for β-galactosidase missing seven N-terminal amino acids and all *lac* regulatory sequences upstream of the *Bam*HI site. Intact *lacY* and *-A* (transacetylase) cistrons are indicated. (C) pMC1403. *lac* fusion vector derivative of pBR322 encoding resistance to ampicillin. (D) pBM1. Insertion of the 5.4-kb viral *Bam*HI fragment into pMC874. The arrow indicates the approximate initiation site and direction of transcription of fusion RNA in *E. coli* cells as determined by S1 mapping. (E) pBM2. Insertion of the 5.4-kb viral *Bam*HI fragment into pMC1403. (F) pBM3. Derived from pBM2 by deletion of viral *Eco*RI fragments. Only part of the LTR downstream of the *Eco*RI site remains. (G) pBM3A. Derived from pBM3 by deletion of the *Sac*I fragment. This deletes most of *lacZ* and the *gag* initiation codon. (H) pBM4. Insertion of the 2.4-kb viral *Eco*RI fragment into pMC1403.

samples were ethanol precipitated and separated on 5% polyacrylamide gels in TBE (90 mM Tris-borate [pH 8.3]–1 mM EDTA)–7 M urea.

Labeling cells with [^{35}S]methionine. Turkey cells infected with PrC virus were labeled with [^{35}S]methionine in Eagle modified medium without methionine at a concentration of 40 μCi/ml (10).

For analysis of fusion proteins, bacterial cells were grown in low-sulfate Tris medium with vitamin B_1 (BLST) medium (5) containing 0.5% glucose and 0.1% Casamino Acids, subcultured in the same medium containing 0.05% Casamino Acids, grown to an absorbance of 1.0 at 600 nm, and labeled by being slowly shaken at 37°C with 0.2 mCi of [^{35}S]methionine per ml.

Wild-type β-galactosidase was prepared from cells containing pBR322 and an F' factor carrying the entire *lac* operon (F'*lac*). F'*lac* extracts were grown by substituting glycerol for glucose and including isopropyl-β-D-thiogalactoside at 1 mM to induce *lac* expression. For analysis of *gag* expression, cells were grown in M63 (19), containing 0.5% glucose and 0.2% Casamino Acids to an absorbance of 0.7 to 1.0 at 600 nm, washed once in M63, and resuspended at the same concentration in methionine assay broth (Difco Laboratories) and 0.2 mCi of [^{35}S]methionine per ml. For pulse-chase experiments, cells were pulse-labeled as described above. After a short time (see below), unlabeled methionine was added at 50 μg/ml. Trichloroacetic acid (TCA) precipitation indicated that this step prevented further incorporation of radioactivity.

Immunoprecipitation. Turkey cell lysates and labeled virions were prepared as described by Conklin et al. (10), except that lysis was sometimes done in the buffer described below for bacterial lysis, a variation which gave less aggregation of virion proteins and more specific immunoprecipitation.

After being labeled, bacterial cells were pelleted, washed once in 10 mM Tris-hydrochloride (pH 7.5), and lysed by being boiled for 2 min in 1% sodium dodecyl sulfate (SDS)–10 mM Tris-hydrochloride (pH 8.0)–1 mM EDTA. Lysates were stored at −20°C. Incorporation was monitored by TCA precipitation.

Immunoprecipitation and SDS-polyacrylamide gel electrophoresis were carried out as described by Eisenman et al. (13). Gels were prepared for fluorography by treatment with En^{3}Hance (New England Nuclear Corp.).

In vitro cleavage of immunoprecipitated Pr76 or fusion proteins with disrupted AMV virions was done as described by Vogt et al. (36). After immunoprecipitation, Staph A pellets (Enzyme Center, Inc.) were suspended in 10 μl of cleavage buffer (50 mM Tris-acetate [pH 7.0]–0.1% Triton–1 M NaCl) and incubated with AMV virions at 37°C.

For preadsorption of anti-*gag* sera, AMV virions, which had been disrupted by being boiled for 2 min in 1% SDS–10 mM Tris (pH 8.0)–1 mM EDTA, were incubated (at 20 μg for each precipitation) with the serum at 4°C for 30 min before the addition of labeled extracts.

Antisera. Sera raised against AMV *gag* proteins were obtained from D. Bolognesi. Anti-β-galactosidase antiserum was a gift of R. C. Huebner (Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Mass.).

Enzymes and reagents. Restriction enzymes, DNA ligase, and polynucleotide kinase were purchased from New England Biolabs, Bethesda Research Laboratories, or New England Nuclear Corp. Calf intestinal alkaline phosphatase and S1 nuclease were from Boehringer-Mannheim. Staph A (IgGsorb) was purchased from the Enzyme Center, Inc.

[^{35}S]methionine (400 Ci/mmol), [γ-^{32}P]ATP, and [α-^{32}P]TTP were from New England Nuclear Corp. or ICN.

RESULTS

Identification of restriction fragments of viral DNA containing bacterial transcriptional and translational initiation sequences. We initially transferred cloned viral DNA to plasmid vectors designed to detect initiation signals for transcription and translation. Plasmids used in these experiments are diagrammed in Fig. 1. The source of viral DNA was the plasmid pATV-8 containing an entire circularly permuted DNA copy of the PrC RSV cloned in pBR322. The plasmid yields infectious virus after transfection of avian cells (P. Norton, unpublished data). The plasmid vector pMC874 (7) contains a fragment of the *E. coli lac* operon encoding an intact *lac* permease (*Y*) gene, and a portion of the β-galactosidase gene (*lac'Z*) extending to a *Bam*HI site near its N terminus. All *lac* operon information upstream of this site is deleted, i.e., seven N-terminal codons (and the initiator AUG), the Shine-Dalgarno sequence for translation of β-galactosidase, and the promoter for transcription of the operon; suitable *E. coli* strains carrying pMC874 thus have a Lac$^-$ phenotype. When a *Bam*HI fragment containing all of these initiation signals is inserted at the *Bam*HI site in pMC874, the resulting plasmid can transform *E. coli* to a Lac$^+$ phenotype, as long as an AUG initiation codon of the insert is in frame with the rest of the β-galactosidase structural gene. To determine whether a portion of RSV DNA could drive *lacZ* expression in *E. coli*, *Bam*HI fragments of pATV-8 were ligated to *Bam*HI-cleaved pMC874, and transformants were plated on media containing the indicator dye XG. A few percent of the transformants were found to have acquired an XG blue phenotype, indicating *lacZ* expression. Analysis of plasmid DNA from several XG blue transformants indicated that all had the 5.4-kb *Bam*HI fragment from pATV-8 fused to the *lac'Z* gene in the orientation expected if the LTR contained a bacterial promoter and if the N-terminal portion of p19^{gag} were supplying an N terminus for β-galactosidase. pBM1 was constructed from one such plasmid (see above).

The region within the 5.4-kb *Bam*HI fragment containing initiation signals was roughly localized by deletion analysis. For this purpose, pBM2 was constructed by transferring the 5.4-kb viral *Bam*HI fragment from pBM1 to a similar vector, pMC1403, which contains the same *lac* sequences as pMC874, except for the absence of an *Eco*RI site in *lac'Z* and the replacement of the *Bam*HI site by an *Eco*RI-*Sma*I-*Bam*HI adapter site (7). Plasmid pBM2 also conferred an XG blue phenotype. Deletion of the left-hand 90% of the cloned fragment from pBM2 by digestion with *Eco*RI and religation to generate pBM3 had no effect on the associated *lac* expression. Thus, the region of viral DNA responsible for the stimulation of *lac* expression was localized to a 585-base pair (bp) region extending from the

*Eco*RI site in U_3 to the *Bam*HI site in *gag*, a region which also contained initiation signals for transcription and translation in eucaryotic cells. The viral origin of the cloned inserts in pBM1, -2, and -3 was confirmed by DNA transfer hybridization with a viral cDNA probe (not shown).

The viral inserts in pBM1, -2, and -3 stimulated similar amounts of *lac'Z* expression as judged by blue color on XG indicator plates and by direct assay of β-galactosidase activity (data not shown). The level of *lac'Z* expression from these high-copy-number plasmids was sufficient to permit growth of the colonies on lactose minimal medium and as red colonies on MacConkey lactose agar (a less sensitive indicator). The inserts also induced expression of the *lacY* (permease) cistron, as indicated by the ability of cells containing these plasmids to grow on plates containing melibiose as a carbon source at 42°C (a Mel^+ phenotype; 19). Since *lacY* contains its own initiation signals for translation, this property allowed us to further delimit the approximate location of the promoter responsible for this expression. pBM3 was digested with *Sac*I, which cleaves once in the leader region preceding *gag* and once within *lacZ*, and ligated to produce pBM3A (Fig. 1). As expected, cells transformed with pBM3A were XG white, yet were still Mel^+. Therefore, the active promoter site must have been within the 308-bp region between the *Eco*RI site in U_3 and the *Sac*I site in the leader region between U_5 and *gag*, i.e., within the same region that contains the probable initiation site of transcription in eucaryotic cells.

Site of initiation of RNA synthesis within viral DNA. The Lac^+ phenotype conferred by pBM1, -2, and -3 should be associated with an mRNA which initiates within the inserted DNA. To confirm this structure, we mapped the 5' end of virus-related transcripts in bacteria containing these plasmids (Fig. 2). A restriction fragment purified from pBM2 extending from the *Xho*I site in *src* to the *Sac*I site in the leader region (Fig. 1) was labeled at its 5' ends and treated with S1 nuclease after hybridization to RNA from bacteria carrying pBM1, -2, or -3. Although both 5' ends of the probe were labeled, we concluded that the end containing the *Sac*I site was protected by these bacterial RNAs, as pBM3 contains no viral sequences upstream of the *Eco*RI site in U_3. The size of the fragment protected by RNA from strains carrying the recombinant plasmids (Fig. 2, lanes c to e) was estimated to be 252 nucleotides, about eight bases smaller than that protected by RNA from quail cells infected with the related viruses NTRE-7 (lane g) or RAV-60 (lane h). The latter fragments had the sizes predicted from the nucleotide sequence of this region (see Fig. 7), with the correction that RNA from eucaryotic cells contains a 5'-capping group which protects an additional 1 to 2 nucleotides from S1 digestion (38). Thus, RNA synthesis from this template in bacteria was initiated about 8 bp downstream from the initiation site in eucaryotic cells.

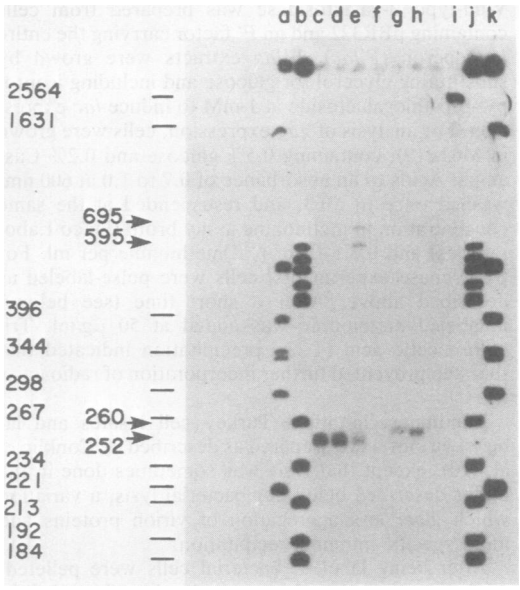

FIG. 2. Analysis of the 5' termini of virus-specific transcripts in *E. coli* and avian cells. A 5' ^{32}P-labeled DNA probe derived from pBM2 and extending from the *Sac*I site in the leader region to the *Xho*I site in *src* was hybridized to 8 μg of RNA from bacteria containing recombinant plasmids or virus-infected avian cells. The resulting hybrids were digested with S1 nuclease, and protected fragments were analyzed by electrophoresis on a 5% polyacrylamide urea gel as described in the text. The RNA was derived from: bacteria containing pBM2 (lane c), pBM3 (lane d), pBM1 (lane e), pMC1403 (lane f), cloned quail cells infected with NTRE-7 (lane g), or RAV-60 (lane h). Marker lanes contained pBR322 cleaved with *Hin*fI (a and k) or *Hae*III (b and j) or probe DNA without digestion with S1 nuclease (i). Arrows indicate sizes of protected fragments. Sizes of marker fragments (in bases) are indicated in margins.

Initiation of protein synthesis within viral RNA. Lac^+ strains carrying pBM1, -2, and -3 should contain a β-galactosidase fusion protein with an N terminus encoded by the viral insert. To examine these proteins, [^{35}S]methionine-labeled cell extracts were immunoprecipitated with antiserum directed against β-galactosidase (Fig. 3, lanes f to j). The β-galactosidase-related proteins encoded by pBM2 and -3 (Fig. 3; lanes h and i) were indistinguishable from authentic β-galactosidase (M_r 116,000) (Fig. 3, lane f) as judged by mobility in SDS-polyacrylamide gels. To verify

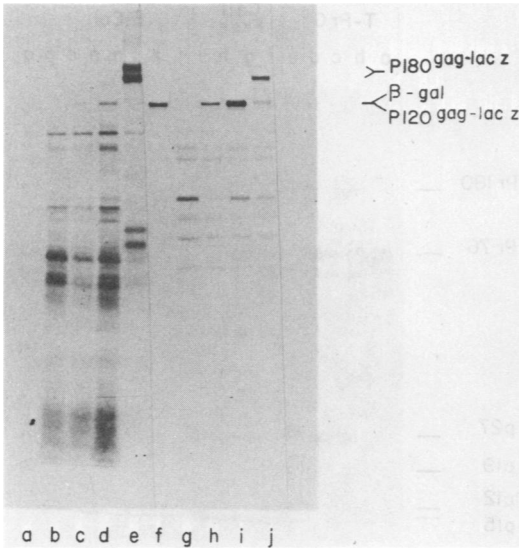

FIG. 3. Immunoprecipitation of *gag*–β-galactosidase fusion proteins. *E. coli* BU8049 cells containing the indicated plasmids were labeled with 0.2 mCi of [^{35}S]methionine per ml in BLST medium for 30 min and lysed; proteins were immunoprecipitated and analyzed on a 12.5% SDS-polyacrylamide gel, as described in the text. Lanes a to e, Immunoprecipitation with antiserum enriched for anti-p19 activity: a, F′*lac* + pBR322; b, pMC1403; c, pBM2; d, pBM3; and e, pBM4. Lanes f to j, Immunoprecipitation with anti-β-galactosidase serum: f, F′*lac* + pBR322; g, pMC1403; h, pBM2; i, pBM3; and j, pBM4. All lysates contained the same amount of TCA-precipitable radioactivity except F′*lac* + pBR322, which contained one-third as much. Two exposure times were used because anti-p19 precipitations were less efficient. P180$^{gag\text{-}lacZ}$ indicates the position of pBM4-specific β-galactosidase-related products. P120$^{gag\text{-}lacZ}$ refers to the fusion protein encoded by pBM2 and pBM3 which is indistinguishable from β-galactosidase (β-gal) on these gels.

that translation initiated in viral information, plasmid pBM4 was constructed (Fig. 1) in which a fragment from the *Eco*RI site in U_3 to one in the p15 portion of *gag* (containing the sequences encoding p19, p27, p12, and part of p15) was fused to *lac′Z*. This was expected to result in a much larger fusion protein, since the nucleotide sequence (25) revealed that this fusion (and the one in pBM3) placed *gag* and *lacZ* in the same reading frame. Cells containing pBM4 had a Lac$^+$ phenotype. Figure 3 (lane j) shows that pBM4 encoded a β-galactosidase-related protein of about M_r 180,000 (P180$^{gag\text{-}lacZ}$), the size predicted for a fusion protein initiating at the N terminus of *gag*. The presence of *gag*-related sequences was confirmed by immunoprecipitation with anti-*gag* serum. Since the inserts in pBM2, pBM3, and pBM4 all contain the N-terminal encoding portion of the *gag* gene, pBM2-, pBM3-, and pBM4-containing cell extracts were treated with antiserum specific for p19 (the N-terminal *gag* peptide), and proteins which comigrated with those precipitated by β-galactosidase antiserum (Fig. 3, lanes c to e) were detected. Bands of similar mobility were not seen with either antiserum when the extract was from cells containing pMC1403 alone (Fig. 3 lanes b and g). Immunoprecipitation of extracts from pBM4-containing cells with either antiserum revealed two plasmid-specific bands (Fig. 3, lane e and j) of about M_r 180,000. The same doublet was also precipitated by monospecific anti-p27 serum (not shown). Anti-β-galactosidase serum in addition precipitated a small amount of a protein from cells containing pBM4 which was about the size of β-galactosidase (Fig. 3, lane j) and had little or no p27 or p19 antigenicity. This pBM4-associated heterogeneity could have been due to cleavage or independent translation initiations (see below). Thus, most or all translation of β-galactosidase was initiated in viral RNA and proceeded in the reading frame of the *gag* gene in pBM2, -3, and -4.

Viral transcription and translation in cells containing pATV-8. Since pATV-8, the source of the fragments which drove *lac′Z* expression in pBM1, -2, -3, and -4, contained the same initiation sequences upstream of a complete *gag* gene, we examined viral transcription and translation in *E. coli* cells containing pATV-8 to determine the state of *gag* gene expression. Estimates of steady-state RNA levels by dot hybridization indicated that transcripts of sequences downstream from U_3 were present at the same level in cells containing pATV-8 as in cells containing the fusion plasmids (about 100 transcripts per cell; data not shown). Preliminary experiments to detect *gag* expression in minicells containing pATV-8, however, showed no protein of the size of Pr76, although the previous results suggested that a protein similar to the *gag* precursor should have been synthesized in *E. coli* containing pATV-8. Suspecting (perhaps incorrectly) that Pr76 was being recognized as foreign and rapidly degraded by *E. coli* proteases, we introduced pATV-8 by transformation into BU8049, a *deg* mutant strain of *E. coli*, because foreign proteins are more slowly degraded in this host (6). We examined *gag* gene expression in these cells by immune precipitation.

Figure 4A shows an analysis of *gag*-related proteins from BU8049 cells containing pATV-8. When cells were labeled for 30 min, Pr76 was again not detected (Fig. 4A, lanes j and k; Fig. 5, lanes a and b); however, prominent bands resembling viral p27 and p15, cleavage products of

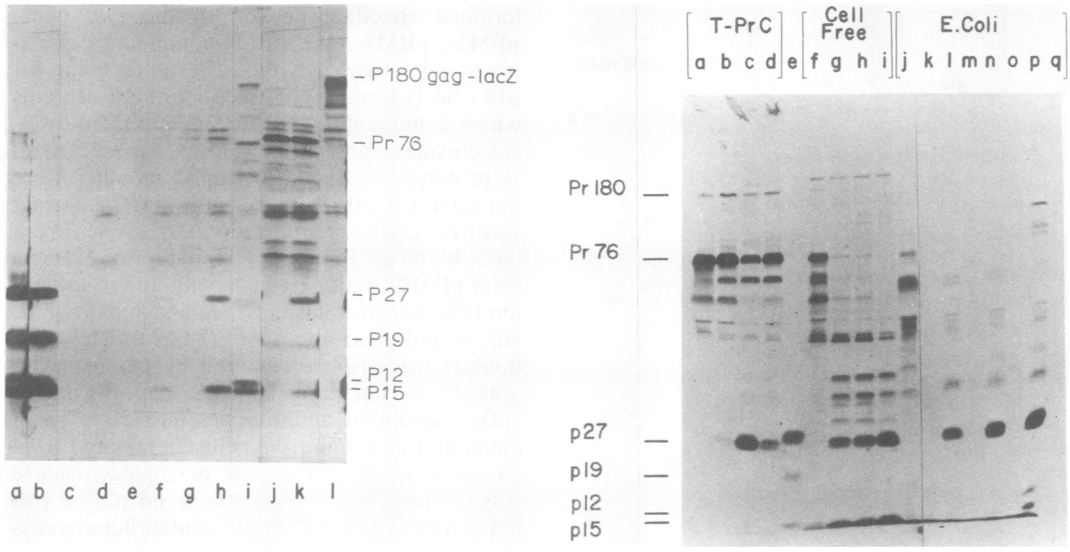

FIG. 4. Synthesis and processing of *gag* proteins in *E. coli* and turkey cells. Bacteria or turkey cells infected with PrC-RSV were labeled with [^{35}S]methionine, and lysates were prepared, immunoprecipitated, and analyzed by electrophoresis on a 12.5% SDS-polyacrylamide gel as described in the text. (A) Immunoprecipitation was with a mixture of antisera specific for the *gag* proteins p15, p19, and p27. Lane a, Purified PrC virions (unprecipitated); lane b, PrC virions immunoprecipitated; lanes c, e and g, control BU8049 cells were pulse-labeled with 0.1 mCi of [^{35}S]methionine for 5 min and chased for 0 min (lane c), 15 min (lane e), or 30 min (lane g); lanes d, f and h, BU8049 cells containing pATV-8 were treated as in lanes c, e, and g; lane i, lysates of PrC-RSV-infected turkey cells labeled for 60 min; lane j, BU8049 cells, 30 min of label; lane k, BU8049 containing pATV-8, 30 min of label; and lane l, RV200 containing pBM4 (the same pBM4-specific proteins were detected when BU8049 cells were used). The positions of relevant proteins are indicated in the margin. Note that p12 migrates more slowly than p15. Lysates for lanes c to h contained 10^6 TCA-precipitable cpm. Those for lanes j to l contained 4×10^6 cpm. (B) Immunoprecipitation with an anti-*gag* antiserum (largely p27 specific). T-PrC: PrC-RSV-infected turkey cells were pulse-labeled for 10 min with 0.2 mCi of [^{35}S]methionine (lane a) and chased for 15 min (lane b), 30 min (lane c), and 60 min (lane d). Lysates containing 10^6 TCA-insoluble cpm were immunoprecipitated for each lane. Lane e, Immunoprecipitated purified PrC virions. Cell-free extract: PrC-infected turkey cells were pulse-labeled for 10 min with 0.2 mCi of [^{35}S]methionine, lysed, and immunoprecipitated (10^6 cpm per lane), and the immunoprecipitate was treated with 5 μg of AMV virions as described in the text for 0 min (lane f), 10 min (lane g), 20 min (lane h), and 60 min (lane i). *E. coli*: BU8049 cells containing pATV-8 were pulse-labeled with 0.2 mCi of [^{35}S]methionine for 2 min (lane j) and chased for 5 min (l), 10 min (n), and 30 min (p). Control BU8049 cells containing pBR322 were treated the same way (lanes k, m, o, and q). Samples of 2×10^6 TCA-insoluble cpm per lane j to q were immunoprecipitated. Sizes of relevant *gag* proteins are indicated.

Pr76, were seen (Fig. 4A, lane k). In the case of p27, we noted that the related peptide synthesized in *E. coli* migrated slightly faster than p27 from virions, but had about the same mobility as p27 in infected turkey cells (Fig. 4A, lane i). This difference in mobility does not appear to be due to phosphorylation (data not shown). We verified that the 27-kilodalton (K) peptide contained p27 sequences, since it was the only *gag*-related peptide precipitated by monospecific anti-p27 serum. We also confirmed that the 15K peptide contained p15 sequences (data not shown). Neither of these peptides was detected when the extract was prepared from cells containing pBM4 (Fig. 4A, lane l) or control BU8049 cells (Fig. 4A, lane j).

After a 5-min pulse-labeling of BU8049 cells containing pATV8 and immunoprecipitation with an anti-*gag* serum, *gag*-specific proteins were only barely visible, but after additional incubation in medium containing unlabeled methionine, 27K and 15K *gag* peptides were detected (Fig. 4A, lanes f and h). These peptides were presumably derived from a heterogeneous set of cleavage intermediates. To confirm this, and since the antiserum used here also precipitated some high-molecular-weight products from the control BU8049 cells (Fig. 4A, lanes c, e, g, and j), the pulse-chase experiment was repeated with another *gag* antiserum which gave a lower background (but precipitated very little p15) and compared with a pulse-chase in turkey cells infected with PrC virus (which was obtained from turkey cells transfected with pATV-8; see above). Figure 4B, lane j, shows proteins from extracts of cells containing pATV8 which

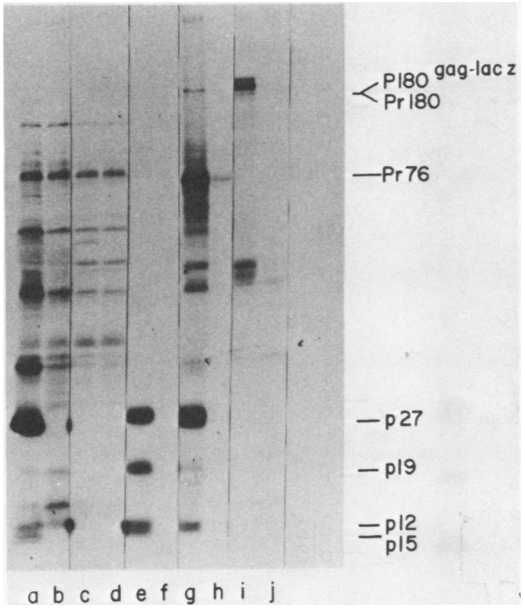

FIG. 5. Verification that the *gag* gene is expressed in *E. coli*. Lysates were precipitated with anti-*gag* sera (a mixture of sera specific for each of the four *gag* peptides) either with or without preabsorption with 20 μg of total AMV virions as described in the text. Lysates were derived from the following cells and contained TCA-insoluble radioactivity as indicated: BU8049 containing pATV-8 (10^7 cpm), lane a, no preabsorption and lane b, + AMV; BU8049 containing pBR322 (10^7 cpm), lane c, no preabsorption and lane d, + AMV; purified PrC virions, lane e, no preabsorption and lane f, + AMV; PrC RSV-infected turkey cells (3×10^6 cpm), lane g, no preabsorption and lane h, + AMV; and BU8049 containing pBM4 (3×10^6 cpm), lane i, no preabsorption and lane j, + AMV.

were pulse-labeled for 2 min and precipitated with this *gag* antiserum. A protein with the mobility of Pr76 was apparent in this sample and was lost within 5 min after the addition of unlabeled methionine (lane l); 27- and 15K peptides were then detectable (lanes l, n, and p). Comparison of the cleavage pattern in *E. coli* with that in turkey cells (lanes a to d) indicated that Pr76 was processed similarly but not identically in the two cell types and much more rapidly in the BU8049 strain.

Figure 4B also shows the products obtained by incubating Pr76 immunoprecipitated from infected turkey cells with a virion extract containing p15 (lanes f to i), since the processing of Pr76 in avian cells is thought to involve the specific proteolytic activity of p15 (11). The kinetics and course of cleavage were similar to those occurring in *E. coli* containing pATV-8 (Fig. 4B, lanes j, l, n, and p). In both cases, the absence of p19 and the presence of a complex set of bands between M_r 30,000 and 76,000 should be noted.

As an additional control, we performed immunoprecipitation with serum that had been preadsorbed with disrupted AMV. The results of this experiment indicated that preabsorption with AMV completely eliminated precipitation of the 27- and 15K peptides from labeled bacterial extracts as well as higher-molecular-weight bands, which were presumably cleavage intermediates (Fig. 5, lanes a and b). Immunoprecipitation of the *gag-lacZ* 180K fusion protein encoded by pBM4 with anti-*gag* serum was also prevented by pretreatment with AMV (Fig. 5, lane i and j). The weak precipitation of a protein of about M_r 60,000 from control cells was also prevented by pretreatment with AMV, possibly reflecting a slight cross-reactivity.

Role of p15 in *gag* processing in *E. coli*. The results presented above indicate that this system should permit the manipulation and analysis of viral gene expression in *E. coli*. In pBM4, for example (Fig. 1), the *gag-lac'Z* fusion may be viewed as resulting in a mutation in p15 deleting the carboxy-terminal 55 codons. pBM4 could thus be used to study the role of p15 in the processing of Pr76 in *E. coli*. The experiment shown in Fig. 4A (lane l) indicates that p27 was not cleaved from the $P180^{gag\text{-}lacZ}$ fusion encoded by pBM4, in spite of the fact that the p27 sequences were present in the *gag* portion of this fusion, just as they were in Pr76 which was cleaved in *E. coli* cells containing pATV-8. The stability of $P180^{gag\text{-}lacZ}$ could be due to the fact that the fusion interrupts p15, and thus there is no p15 in pBM4-containing cells. Alternatively, the conformation of the *gag* sequences in the $P180^{gag\text{-}lacZ}$ fusion protein could be such that p27 is not susceptible to cleavage by an *E. coli* protease. The experiment shown in Fig. 6 was designed to determine whether the absence of p15 could account for this stability. The addition of a virus extract containing p15 to an immunoprecipitate of $P180^{gag\text{-}lacZ}$ (Fig. 6, lane j) resulted in the loss of the 180K bands and the appearance of a band with the mobility of the p27 product seen in *E. coli* cells containing pATV-8 (Fig. 6, lanes k and l). To determine whether this 27K cleavage product was derived from *gag* sequences, the in vitro cleavage was repeated with extracts from cells containing pBM4, pBM3, and authentic β-galactosidase which had been precipitated with β-galactosidase antiserum. The result indicated that β-galactosidase was cleaved by p15 under these conditions (Fig. 6, lanes a to c). The fusion protein encoded by pBM3 was cleaved in a similar fashion (Fig. 6, lanes d to f). In both cases, however, although a large β-galactosidase cleavage product (indicated by an arrow in Fig. 6) was apparent, no 27K

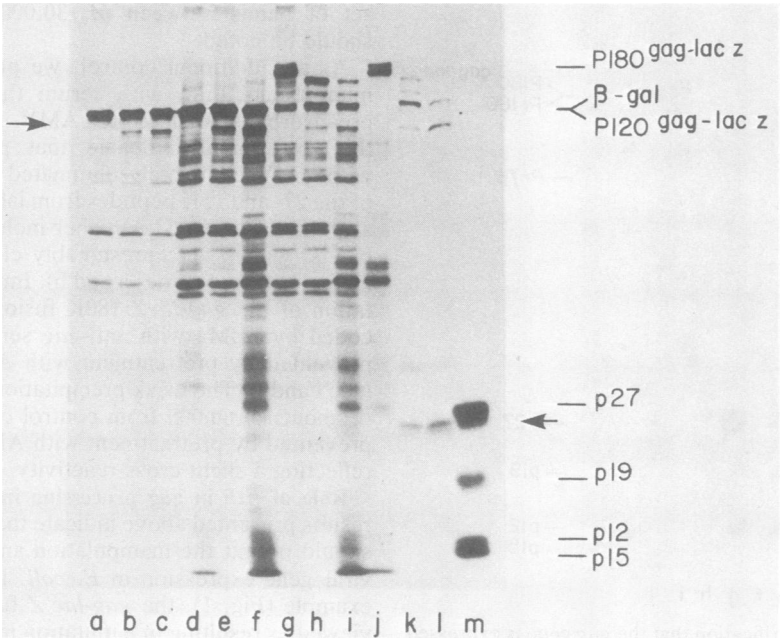

FIG. 6. In vitro cleavage of P180$^{gag\text{-}lacZ}$ by p15. [^{35}S]methionine-labeled lysates of BU8049 *E. coli* were prepared, immunoprecipitated with anti-β-galactosidase serum (lanes a to i) or anti-p19 serum (lanes j to l), and resuspended in AMV cleavage buffer as described in the text. They were then incubated at 37°C for 1 h with or without AMV virions as indicated, before loading on a 12.5% SDS-polyacrylamide gel. F'*lac* + pBR322: No AMV (lane a), 5 μg of AMV (lane b), and 20 μg of AMV (lane c). pBM3: No AMV (lane d), 5 μg of AMV (lane e), and 20 μg of AMV (lane f). pBM4: No AMV (lane g), 5 μg of AMV (lane h), and 20 μg of AMV (lane i). pBM4: No AMV (lane j), 5 μg of AMV (lane k), and 20 μg of AMV (lane l). Lane m contained labeled PrC virions. All lysates contained the same amount of TCA-precipitable radioactivity except F'*lac* + pBR322 (lanes a to c), which contained one-third as much. The right-pointing arrow indicates the position of a β-galactosidase-specific cleavage product. The left-pointing arrow indicates the position of a pBM4-specific cleavage product.

product was generated. Cleavage of nonviral proteins by p15 has been previously reported (12). Cleavage of P180$^{gag\text{-}lacZ}$ which had been precipitated with this serum again resulted in a 27K product (Fig. 6, lanes g to i), although the cleavage was much less efficient relative to cleavage after precipitation with p19 antiserum (Fig. 6, lanes j to l). These results indicated that the appearance of the 27K peptide after in vitro cleavage was pBM4 specific. Although a total AMV virion preparation, rather than purified p15, was used for these experiments, the appearance of a 27K protein as a specific cleavage product of P180$^{gag\text{-}lacZ}$, and the concordance of the cleavage pattern of Pr76 (Fig. 4B, lanes f to i) with that reported for p15 (36), argue that possible contaminating proteases were not involved.

DISCUSSION

The use of *E. coli* genetic fusion vectors for manipulation of avian retroviruses. To take full advantage of *E. coli* genetic fusion vectors in the study of foreign genes, it is desirable that transcription and translation initiate within the inserted DNA. In our case, since the inserts were derived from an avian virus, we felt that it was particularly important to verify that this was the case and to determine which viral sequences were involved. It had previously been shown that *E. coli* RNA polymerase holoenzyme bound to the LTR in vitro (14), and that the LTR had a promoter activity in *E. coli* (20). We determined that this activity accounted for the *lac* operon expression in our fusion plasmids, that *gag* peptide sequences provided an N terminus for active β-galactosidase, and that the sequences involved should be subject to manipulation in ways that would alter sequences involved in viral expression.

The viral LTR contains a bacterial promoter. Salient features of the nucleotide sequence of the RSV-*lacZ* fusion region as determined by analysis of pBM3 are shown in Fig. 7. In plasmids with LTR sequences fused to the structural gene for β-galactosidase, we have localized a bacterial promoter sequence within the U_3 region. The nucleotide sequence shown in Fig. 7 is numbered as a conventional proviral sequence

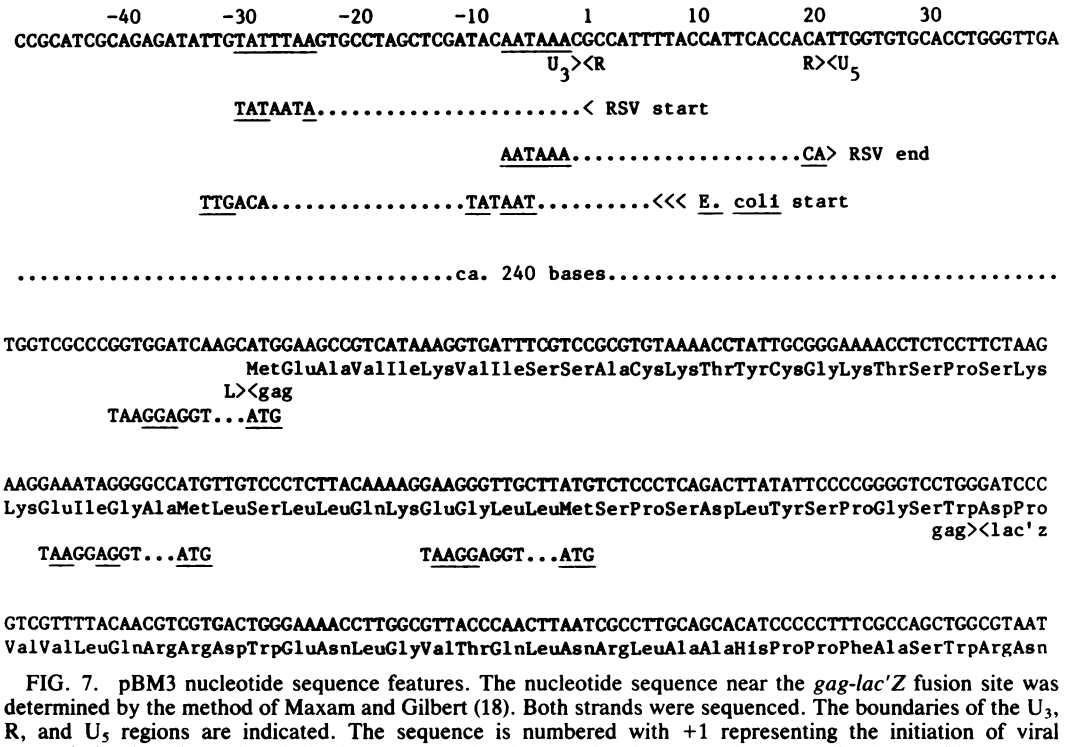

FIG. 7. pBM3 nucleotide sequence features. The nucleotide sequence near the *gag-lac'Z* fusion site was determined by the method of Maxam and Gilbert (18). Both strands were sequenced. The boundaries of the U_3, R, and U_5 regions are indicated. The sequence is numbered with +1 representing the initiation of viral transcription in avian cells within the 5' LTR, with the associated Hogness box underlined at position −30 to −23. The polyadenylation signal recognized at the 3' LTR at −8 to −1 is also underlined, and the polyadenylation site at the end of R is indicated as RSV end. The start site for transcription in *E. coli* and the promoter sequences recognized in *E. coli* as the Pribnow box and −35 sequence are shown, with the extent of homology with the consensus sequences underlined. With respect to translation, the three methionine codons in the p19 frame are delineated with the extent of homology upstream from each, with the consensus Shine-Dalgarno sequence indicated by a line under the consensus sequence.

with the initiation site for eucaryotic transcription at the beginning of R at number 1. A typical Hogness box is located at position −23. The sequence at the end of U_3 (AATAAA) signals polyadenylation at the end of the 3' R sequence. Based on the results shown in Fig. 2, we have located the initiation site for viral transcription in *E. coli* within a few nucleotides of nucleotide 9 (an A residue) (Fig. 7). Taking this as the initiation site, and using the structure of the consensus bacterial promoter as a model (28), we would place the Pribnow box at −1 and the −35 sequence at −28. The sequences found at these positions in the RSV genome are similar to the *E. coli* promoter consensus sequence: TACAAT compared with TATAAT for the Pribnow box and TTGTAT versus TTGACA for the −35 sequence. Thus, the Pribnow box apparently recognized in *E. coli* overlaps with the viral polyadenylation signal (AATAAA at position −8 to −2), and the −35 sequence would overlap the viral Hogness box (−30 to −23). This assignment confirms the recently proposed location of this promoter obtained by deletion mapping (20).

Translation initiation in *E. coli* within viral DNA. Figure 7 also diagrams sequences around the site of initiation of viral *gag* protein synthesis in RSV-infected avian cells. The initiator codon AUG is separated from the 5' end of the mRNA by a nontranslated leader sequence of about 250 nucleotides. Although we have not precisely located the initiation site for β-galactosidase translation in *E. coli* containing pBM1, -2, -3, and -4, our results indicate that translation of the fusion proteins initiates in the *gag* frame within viral information. Inspection of the *gag* nucleotide sequences present in pBM3 (Fig. 7) and pBM4 (25) indicates that *gag* and β-galactosidase are translated in the same reading frame in these fusions. We can also infer translation in the *gag* reading frame from the size of the $P180^{gag\text{-}lacZ}$ fusion in pBM4-containing cells and the presence of both *gag* and β-galactosidase antigens on these proteins (Fig. 3). In pBM3, there are three methionine codons preceding the *Bam*HI site at the p19-*lacZ* border in the *gag* (p19) frame. Each is preceded by a partial homology with the consensus Shine-Dalgarno ribo-

some binding sequence (27). The upstream AUG is the initiation site for *gag* synthesis in avian cells (11). We cannot at present determine which of these sites is used in bacteria.

We do not know whether the same initiation site for translation as is used in pBM3 is used exclusively in pBM4. Immunoprecipitation of extracts from pBM4-containing cells reveals some heterogeneity in the size of *gag*-containing fusion proteins (Fig. 3, lanes e and j). pBM4-containing cell extracts also contain a β-galactosidase-related protein which is indistinguishable from β-galactosidase on the basis of size and which contains no detectable *gag* determinants (Fig. 3, lane j). This protein probably does not represent another initiation site for translation near the fusion site since a frameshift mutation at the *Bgl*II site within the p27 portion of pBM4 results in an XG white, Mel^+ phenotype (data not shown). This band and the presence of a doublet near 180K (Fig. 3, lanes e and j) may be due to cleavage within the *gag* domain by an *E. coli* protease or could reflect residual proteolytic activity from the fragment of p15 retained in pBM4, which includes the N-terminal 69 of the 124 p15 amino acids.

***gag* gene expression in *E. coli*.** We found that the *gag* gene is expressed in *E. coli* cells containing pATV-8, presumably from the same initiation sites driving expression of the *gag*–β-galactosidase fusion proteins in pBM1, -2, -3, and -4. The predominant stable *gag*-related proteins in these cells are *gag* cleavage products rather than the primary polyprotein Pr76 (Fig. 5, lanes a and b). This results from posttranslational processing of Pr76 (Fig. 4). The half-life of Pr76 in *E. coli*, under the conditions used here, is less than 5 min, in contrast with a 45-min half-life in RSV-infected turkey cells (Fig. 4B and reference 11). Although we do not know the basis for this difference, it may be related to the unexplained variation in the stability of Pr76 synthesized in vitro with eucaryotic or bacterial cell extracts. (23, 29). The kinetics of processing of Pr76 could be a function of several parameters, including intracellular concentration, host factors, and subcellular location. We recently found that the conditions of growth and labeling of the cells may also affect the kinetics of cleavage of Pr76 in *E. coli*. The course of Pr76 cleavage in *E. coli* closely resembles that occurring in turkey cells, with respect to intermediates and stable products. We verified that stable peptides with the mobilities and antigenicities of p27 and p15 are produced. We have been unable to determine the status of p19 and p12 in *E. coli* containing pATV-8. In the latter case, we did not have an antiserum which was capable of immunoprecipitating p12 in large amounts. In the case of p19, a combination of several factors may have contributed to our inability to detect a counterpart in *E. coli*. The cleavage of Pr76 in *E. coli* may resemble that occurring in vitro upon addition of p15 to the extent that other host factors and proteases involved in Pr76 processing in avian cells would be absent. p19 is not seen after in vitro cleavage of Pr76 (Fig. 4B, and reference 13). Alternatively, the size of p19 or a partial cleavage product containing p19 would depend upon the intiation site for translation, whose exact position we were unable to determine. Finally, a portion of p19 may be a substrate for a bacterial protease.

The most likely explanation for the processing of Pr76 synthesized in *E. coli* was that p15 was involved. That this was the case was suggested by comparison of the course of cleavage in *E. coli* with that occurring in vitro upon addition of p15 (Fig. 4B). We were able to test this hypothesis genetically by using the *gag*–β-galactosidase fusion protein encoded by pBM4 which contains all of the *gag* sequences except the carboxy-terminal 55 amino acids of p15 downstream of the *Eco*RI site fusing *gag* sequences to *lacZ*. This protein was relatively stable, and no *gag* cleavage products were seen (Fig. 4A, lane l). Since this fusion protein was a substrate for cleavage by purified AMV virions in vitro (Fig. 6), we concluded that the reason for its stability in vivo relative to Pr76 was the absence of functional p15. We intend to establish this point more completely by adding a plasmid to pBM4-containing cells which encodes p15. Whereas these results strongly implicate p15 as the enzyme responsible for much of the cleavage of Pr76 in *E. coli*, they do not bear on the question as to whether a host enzyme is involved in the initial cleavage of p15 from Pr76 or whether p15 is active as part of the precursor (11). While examining the susceptibility of $P180^{gag\text{-}lacZ}$ to in vitro cleavage by p15, we were surprised to find that the efficiency of cleavage depended upon the antiserum used (Fig. 6). A similar effect on cleavage of Pr76 had been previously reported when different anti-*gag* sera were used (21).

It is probably fortuitous that viral DNA contains sequences recognized in *E. coli* for initiation of transcription and translation and that in our case these signals are at or near those which are active in RSV-infected avian cells. In Fig. 2, lane e, for example, minor transcription initiation sites in pBM1, upstream of the one described above (and upstream of the *Eco*RI site in U_3), can be detected. We observed, in fact, that a fragment containing sequences upstream of the *Eco*RI site in U_3, when fused to *lacZ*, can drive expression of β-galactosidase to the extent that colonies containing this fusion were blue on XG plates, but such cells contained no assayable β-galactosidase activity in vitro and no immuno-

precipitable fusion proteins were detectable on gels (not shown). In addition, promoter activity has been observed on the nonsense strand of restriction fragments from both cloned Abelson murine leukemia virus and mink cell focus-forming virus (R. C. Huebner and J. P. Stoye, personal communication).

Taken together, these results indicate that the cloning of retroviral genes in *E. coli* should be a useful procedure for the manipulation of the viral genome. The *gag-lacZ* fusion plasmids should permit the construction and selection of point mutations in viral DNA by standard bacterial genetic techniques.

The finding that these viral sequences drive β-galactosidase expression in *E. coli* makes it quite probable that these constructions would permit β-galactosidase expression in avian cells. This might permit not only the quantitative study of viral promoter function in these cells, but also serve as a model system to assess the suitability of this and other constructions for the insertion of other genes into the retroviral genome. Recently, plasmids derived from those discussed here have been introduced into avian cells, and similar fusion proteins were expressed and detectable without selection (P. Norton, unpublished data). In addition, the fusion RNA was packaged into infectious virions in the presence of helper virus. The latter result was not unexpected as the structures of these *gag-lacZ* genetic fusions are analogous to those which fuse *gag* sequences to *onc* genes in the defective transforming viruses (8). It is thus possible that these transforming proteins could be produced in their native form simply by introducing the relevant viral DNA into *E. coli* or manipulated in *E. coli* by procedures similar to the ones outlined here.

ACKNOWLEDGMENTS

We thank M. Champion for the nucleotide sequence analyses and J. Stoye and P. Norton for helpful discussion. B.M. thanks Glenn Albrecht, Kathy Conklin, Steve Herman, Robert Cooney, and Mary Bostic-Fitzgerald.

This work was supported by Public Health Service grants RO1-CA-27108 from the National Cancer Institute (to J.M.C.) and RO1-A1-15840 from the National Institute of Allergy and Infectious Disease (to M.M.). J.M.C. was the recipient of a Faculty Research Award from the American Cancer Society, and B.M. was supported in part by training grant GM 17310 from the National Institutes of Health.

LITERATURE CITED

1. **Beckwith, J.** 1981. A genetic approach to characterizing complex promoters in *E. coli*. Cell **23**:307–308.
2. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenoviral mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
3. **Bishop, J. M., S. A. Courtneidge, A. D. Levinson, H. Oppermann, N. Quintrell, D. K. Sheiness, S. R. Weiss, and H. E. Varmus.** 1980. The origin and function of avian retrovirus transforming genes. Cold Spring Harbor Symp. Quant. Biol. **44**:919–930.
4. **Blair, D. G., M. Oskarsson, T. G. Wood, W. L. McClements, P. J. Fischinger, and G. F. Vande Woude.** 1981. Activation of the transforming potential of a normal cell sequence: a molecular model for oncogenesis. Science **212**:941–943.
5. **Blumberg, D. D., and M. H. Malamy.** 1974. Evidence for the presence of nontranslated T7 late mRNA in infected F′(PIF⁻) episome-containing cells. J. Virol. **13**:378–385.
6. **Bukhari, A., and D. Zipser.** 1973. Mutants of *Escherichia coli* with a defect in the degradation of nonsense fragments. Nature (London) New Biol. **243**:238–241.
7. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
8. **Coffin, J. M.** 1982. Structure of the retroviral genome, p. 261–368. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), Molecular biology of tumor viruses. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Coffin, J. M., P. N. Tsichlis, C. S. Barker, S. Voynow, and H. L. Robinson.** 1980. Variation in avian retrovirus genomes. Ann. N.Y. Acad. Sci. **354**:410–425.
10. **Conklin, K. F., J. M. Coffin, H. L. Robinson, M. Groudine, and R. Eisenman.** 1982. Role of methylation in the induced and spontaneous expression of the avian endogenous virus *ev*-1: DNA structure and gene products. Mol. Cell. Biol. **2**:638–652.
11. **Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich.** 1982. Protein biosynthesis and assembly, p. 513–648. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Dittmar, K. J., and K. Moelling.** 1978. Biochemical properties of p15-associated protease in an avian RNA tumor virus. J. Virol. **28**:106–118.
13. **Eisenman, R. N., W. S. Mason, and M. Linial.** 1980. Synthesis and processing of polymerase proteins of wild-type and mutant avian retroviruses. J. Virol. **36**:62–78.
14. **Guntaka, R. V., P. Y. Rao, R. A. Katz, and S. A. Mitsialis.** 1980. Binding of *E. coli* RNA polymerase to a specific site located near the 3′ end of the avian sarcoma virus genome. Biochim. Biophys. Acta **607**:457–469.
15. **Hsu, T. W., J. L. Sabran, G. E. Mark, R. V. Guntaka, and J. M. Taylor.** 1978. Analysis of unintegrated avian RNA tumor virus double-stranded DNA intermediates. J. Virol. **28**:810–818.
16. **Katz, R. A., C. A. Omer, J. H. Weis, S. A. Mitsialis, A. J. Faras, and R. V. Guntaka.** 1982. Restriction endonuclease and nucleotide sequence analyses of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions. J. Virol. **42**:346–351.
17. **Leis, J., J. McGinnis, and R. W. Green.** 1978. Rous sarcoma virus p19 binds to specific double-stranded regions of viral RNA: effects of p10 on cleavage of viral RNA by RNase III. Virology **84**:87–98.
18. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
19. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Mitsialis, S. A., J. F. Young, P. Palese, and R. Guntaka.** 1981. An avian tumor virus promoter directs expression of plasmid genes in *E. coli*. Gene **16**:217–225.
21. **Moelling, K., A. Scott, K. E. J. Dittmar, and M. Owada.** 1980. Effect of p15-associated protease from an avian RNA tumor virus on avian virus-specific polyprotein precursors. J. Virol. **33**:680–688.
22. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell **23**:323–334.
23. **Pawson, T., G. S. Martin, and A. E. Smith.** 1976. Cell-free

translation of virion RNA from nondefective and transformation-defective Rous sarcoma virus. J. Virol. **19**:950–967.

24. **Payne, G. S., S. A. Courtneidge, L. B. Crittenden, A. M. Fadly, J. M. Bishop, and H. E. Varmus.** 1981. Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell **23**:311–322.
25. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
26. **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 forms of avian sarcoma virus DNA: both termini of linear DNA bear a 300 nucleotide sequence present once or twice in two species of circular DNA. Cell **15**:1383–1395.
27. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. U.S.A. **71**:1342–1346.
28. **Siebenlist, U., R. B. Simpson, and W. Gilbert.** 1980. *E. coli* RNA polymerase interacts homologously with two different promoters. Cell **20**:269–281.
29. **Siegert, W., R. N. H. Konings, H. Bauer, and P. H. Hofschneider.** 1972. Translation of avian myeloblastosis virus RNA in a cell-free lysate of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **69**:888–891.
30. **Temin, H. M., and D. Baltimore.** 1972. RNA-directed DNA synthesis and RNA tumor viruses. Adv. Virus Res. **17**:129–186.
31. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. **77**:5201–5205.
32. **Tsichlis, P. N., and J. M. Coffin.** 1980. Recombinants between endogenous and exogenous avian tumor viruses: role of the C region and other portions of the genome in the control of replication and transformation. J. Virol. **33**:238–249.
33. **Tsichlis, P. N., and J. M. Coffin.** 1980. Role of the C region in relative growth rates of endogenous and exogenous avian oncoviruses. Cold Spring Harbor Symp. Quant. Biol. **44**:1123–1132.
34. **Varmus, H. E.** 1982. Replication of retroviruses, p. 369–513. *In* R. Weiss, N. Teich, H. Varmus, J. Coffin (ed.), Molecular biology of tumor viruses. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Vogt, P. K.** 1976. Genetics of RNA tumor viruses, p. 341–455. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 9, Plenum Publishing Corp., New York.
36. **Vogt, V. M., A. Wight, and R. Eisenman.** 1979. *In vitro* cleavage of avian retrovirus *gag* proteins by viral protease, p15. Virology **98**:154–167.
37. **Von der Helm, K., S. Kemperi, W. Willie, and K. Willecke.** 1980. Rous sarcoma virus precursor protein Pr76 is processed in avian sarcoma virus-transformed mammalian cells after fusion-injection of viral protein p15. Virology **106**:310–316.
38. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin m-RNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.