The Human Immunodeficiency Virus Type 1 gag Gene Encodes an Internal Ribosome Entry Site

CHRISTOPHER B. BUCK,1 XUEFEI SHEN,2 MICHAEL A. EGAN,2 THEODORE C. PIERSON,3
CHRISTOPHER M. WALKER,4 AND ROBERT F. SILICIANO3*

Program in Cellular and Molecular Medicine,1 Program in Biochemistry, Cellular and Molecular Medicine,2 and Graduate Program in Immunology,3 Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, and Department of Virology, Chiron Corporation, Emeryville, California 946084

Received 8 May 2000/Accepted 5 October 2000

Several retroviruses have recently been shown to promote translation of their gag gene products by internal ribosome entry. In this report, we show that mRNAs containing the human immunodeficiency virus type 1 (HIV-1) gag open reading frame (ORF) exhibit internal ribosome entry site (IRES) activity that can promote translational initiation of Pr55gag. Remarkably, this IRES activity is driven by sequences within the gag ORF itself and is not dependent on the native gag 5′-untranslated region (UTR). This cap-independent mechanism for Pr55gag translation may help explain the high levels of translation of this protein in the face of major RNA structural barriers to scanning ribosomes found in the gag 5′ UTR. The gag IRES activity described here also drives translation of a novel 40-kDa Gag isoform through translational initiation at an internal AUG codon found near the amino terminus of the Pr55gag capsid domain. Our findings suggest that this low-abundance Gag isoform may be important for wild-type replication of HIV-1 in cultured cells. The activities of the HIV-1 gag IRES may be an important feature of the HIV-1 life cycle and could serve as a novel target for antiretroviral therapeutic strategies.

Like other complex retroviruses, human immunodeficiency virus type 1 (HIV-1) employs a variety of mechanisms to express numerous proteins from a single-genomic-length RNA transcript. One such mechanism is the regulated splicing of the primary transcript into more than 30 distinct mRNA species (65, 74–76). Other mechanisms for the production of different polypeptides from mature viral mRNA species rely on translational events that are atypical of normal host protein translation.

The initiation of translation of most eukaryotic mRNAs is thought to occur by a process involving ribosome scanning. In this model, the 40S ribosomal subunit binds to the 5′ end of a capped mRNA (80, 81), commences scanning toward the 3′ end, and initiates translation upon encountering an AUG codon in suitable context (often referred to as a “Kozak” context) (47, 49, 50). The 60S ribosomal subunit is then recruited, and translation of a single polypeptide begins.

Translation of some HIV-1 mRNAs can depart from this basic model in several ways. Translation of the pol open reading frame (ORF) requires a −1 frameshift at a “slippery” sequence within the gag ORF. This frameshift allows roughly 5 to 10% of translating ribosomes to bypass the normal gag stop codon, enabling translation of a Gag-Pol precursor protein (reviewed in reference 32). Other HIV-1 gene products, including Env and Nef, have been shown to be translated by “leaky” scanning of ribosomes past the vpu and rev AUGs, respectively (74, 75, 77).

Another alternative mechanism for translational initiation is driven by RNA structural elements referred to as internal ribosome entry sites (or segments) (IRESs). IRESs are thought to promote initiation of translation by directly binding ribosomes, with participation of other host cell factors, in a manner independent of the mRNA cap or of scanning through upstream sequences. Since the initial discovery of an IRES in the 5′ untranslated region (UTR) of picornaviruses (43, 66), IRES activity has been described in a wide variety of viral and cellular mRNAs (reviewed in references 41 and 71). Details of the molecular mechanisms through which these RNA structures promote the internal entry of ribosomes are the subject of much current interest.

IRESs have been described in several retroviruses (4, 5, 19, 63, 91), leading to speculation that HIV-1 mRNAs might exhibit IRES activity. The data presented in this report indicate that mRNAs containing the HIV-1 gag ORF exhibit IRES activity that can drive production of both the familiar Gag precursor protein, Pr55gag, and a novel 40-kDa N-terminally truncated isoform of Gag (p40). We demonstrate that this novel Gag isoform is present in cultured cells infected with HIV-1. Our findings suggest that expression of p40 may be required for wild-type viral growth kinetics.

Unlike the majority of other IRESs, which have been found partially or wholly within the 5′ UTRs of mRNAs, the IRES activity reported here is driven by sequences entirely within the gag ORF. This surprising arrangement is similar to the IRES of Moloney murine leukemia virus (MMLV) (91). The observation of IRES activity in a variety of retroviral species suggests that such activity may be an important feature of retroviral biology. The HIV-1 gag IRES described here might thus serve as a novel target for viral attenuation or antiretroviral therapeutics.
Vaccinia virus expression vectors. Vaccinia virus vB2 was generated by ligation of a PCR product of the HBX2 gag gene from plasmid pHBX2neo (kindly provided by Xiao Fang Yu, John Hopkins University) into the XhoI-NheI sites of vector pCS1-MCS1 (32). This plasmid, pAGT, 8 nucleotides (nt) downstream of an ATG codon in strong Kozak context (AGT ACC ATG G) found in the multiconing site of this vector, resulting in somewhat diminished Pr55<sup>env</sup> expression compared to unconstructed Pr55<sup>env</sup>. Virus vENyVAG was generated by first transfecting the XhoI-XhoII fragment of plasmid vector pENVyVAG (see below) into pCS1MCS1 XhoI-NheI vector species. The KpiI- EcoRI<sup>I</sup> env fragment of the resulting recombinant plasmid was replaced with the KpiI-EcoRI<sup>I</sup> env fragment of plasmid pPE15 (kindly provided by Patricia Earl, National Institutes of Health) in order to remove the two vaccinia virus early transcription termination signals found in this portion of <i>env</i> (20). The resulting construct, pVEG, encodes a hybrid <i>env</i> ORF derived from two closely related <i>env</i> sequences, HIV-1 LAI (pENV-<i>GAG</i>) and HIV-1 BH5 (pPE15). Recombinant vaccinia viruses were generated by recombination in strain vWVR-Lvar and subsequent plaque purification according to a method described previously (21).

Plasmid expression vectors. Plasmid construction was accomplished by standard methods. All PCR inserts and splice junctions were verified by sequencing. The vectors were based on the Pr55<sup>env</sup> gene of HIV-1 isolate HXB2 (HIV Sequence Database, 1997 [http://hiv-web.lanl.gov/], Los Alamos National Laboratory Theoretical Biology and Biophysics, Los Alamos, NM, The.). The Pr55<sup>env</sup> gene was silently modified to introduce an XhoI site near its 5' end (G<sub>3</sub>ATG<sub>3</sub>T - C<sub>4</sub>ATP<sub>4</sub> - C) to attenuate the slippery sequence which promotes ribosomal frameshifting to the pol reading frame as described earlier (70). The Kozak context surrounding the start codon Pr55<sup>env</sup> in these vectors is CACACCATG G. The HIV-1 <i>env</i> gene was derived from provirus pLAI.

Vectors pEMC-<i>CAT</i>, pCAT, pCAT-GAG, pENVyVAG, pENV-<i>GAG</i>, pENV-<i>ΔUTR-GAG</i>, pENV-<i>ΨGAG</i> and pENV-<i>ΔGAG</i> were all expressed from the backbone of pCI (Promega). Vectors pLC2Z, pIGAG, pΔ40, pΔ55, pFS, and pSFA40 were expressed from the backbone of plasmid pCMV6c55gag-pre, which was kindly provided by Chiron Corp. (Emeryville, Calif.). All vectors contain cytomegalovirus (CMV) promoter and an intron in the 5' UTR to boost expression. All of the vectors also bear the hepatitis B virus (HBV) posttranscriptional regulatory element (PRE) (corresponding to the roughly 570-bp Sp<sup>II</sup>-FspI fragment of HBV strain adw 7) in the 3' UTR to help overcome the Rev-dependency of gag and env mRNAs (35, 36).

Plasmids pGAG and pGAG contain the silently modified HBX2 Pr55<sup>env</sup> gene. pΔ40 bears an ATG<sub>1</sub>-Δ2<sup>I</sup> - Δ3<sup>Ψ</sup>-Δ4<sup>G</sup> mutation generated by PCR-directed mutagenesis. pΔ55 was made by intramolecular ligation of blunted NheI and XhoI sites surrounding ATG<sub>1</sub>, resulting in a 7-nt truncation at the 5' end of Pr55<sup>env</sup>. pFS bears a single G deletion from a string of five G residues at gag bases 28 to 32.

The control vector pCAT was generated by ligation of the KpiI-XhoII chloramphenicol acetyltransferase (CAT) gene fragment from vector pCDNA3.1CAT (Invitrogen) into vector pGAG such that CAT replaced Pr55<sup>env</sup> between the intron and the HBV PRE. Vector pENVyVAG was generated by three-way ligation of a plasmid <i>env</i> gene fragment and a short PCR product into vector pCI (Promega). pENV-<i>ΔUTR-GAG</i> was generated by first transferring the XhoI-XhoII fragment of the resulting construct into vector pCMV6c55gag-pre and then blunting the 5' ends with TfiI and religating, thus removing all gag and pol coding regions.

Northern blotting probe pHBG was generated by transcribing the HindIII-MfeI fragment of gag from vector pGAG into an plbluescript trimmed with EcoRI-HindIII. This vector was linearized with XhoI prior to transcription with T7 RNA polymerase (see Northern blotting section, below). The resulting probe is homologous to bases 923 to 1183 of the Pr55<sup>env</sup> ORF and is also homologous to a 28-bp section of the 3' UTR of all pCI-based vectors (including pCAT).

Large-scale plasmid preps for transfection purposes were made by EndoFree Plasmid Maxi kit (Qiagen). Provirail plasmids. The SapI-SphI fragment of ATG<sub>1</sub>-ΔCAT mutant plasmid pΔ40 was transferred into plasmid-cut pLAI SapI-SphI to create provirus pVATG. Because this SapI-SphI fragment bears a nonsilent S<sub>1</sub>S<sub>2</sub>-<i>Δ</i>AsnHXB2 mutation at codon 126, an isogenic control provirus, pVATG<sub>Δ</sub>, was generated using the SapI-SphI fragment of unmutated pGAG.

Cell culture and transfection. Peripheral blood mononuclear cells (PBMC) were obtained by Ficoll purification of blood from laboratory volunteers, and maintained in STCM (RPMI, 10% fetal calf serum [FCS], interleukin-2 [100 U/ml], T-cell growth factor). PBMC were activated by treatment for 3 days with 0.25 mg of phytomethylglutamid (PHA) per ml. Activated PBMC cultures were depleted by treatment with anti-CD8 magnetic beads (Dynal). Epstein-Barr virus (EBV)-transformed B cells were derived from patient A42 (32). COS-7 and 293T cell lines were obtained from the American Type Culture Collection (ATCC) and maintained according to conditions described in the ATCC catalog.

COS-7 cells were transfected in 12-well plates with Lipofectamine Plus (Life Technologies, Inc.) and were subjected to radioimmunoprecipitation assay (RIPA) or p24 enzyme-linked immunosorbent assay (ELISA) at 24 h posttransfection. For poliovirus infection experiments, 3 × 10<sup>5</sup> 293T cells were preplated overnight in a 25-cm<sup>2</sup> flask and then transfected with Lipofectamine 2000 (Life Technologies). At 24 h posttransfection, cells were divided into eight wells of a 12-well plate (4 cm<sup>2</sup>/well) and then infected with poliovirus beginning 40 h posttransfection. Infectious HIV-1 supernatants were collected from provirus-transfected COS-7 cells 3 days after transfection. The stable Jurkat lines Jurkat + sac and Jurkat△Sac were generated by electroporation of Jurkat cells with 10 μg of BpHI-linearized pSAC or pSAC in RPMI at 270 mV and 960 μF. The transfected cells were incubated for 2 days and then subjected to selection in 350 μg of hygromycin B (Boehringer) per ml for 3 weeks. Cells were then sorted using a fluorescence-activated cell sorter for EGFP expression, and the brightest 10% of the FL<sup>1</sup> cells were collected (not shown). The resulting polyclonal populations were maintained in the presence of 350 μg of hygromycin B per ml.

Immunoprecipitation. Pulse-chase experiments on vaccinia virus-infected cells were performed as described elsewhere (86). Briefly, 1 million A42 EBV cells were vaccinia virus infected for 4 h, starved for 30 min in starvation medium (Cys·Met·Dubcaco modified Eagle medium [DME], 5% dialyzed FCS [Life Technologies]), and then radiolabeled for 30 min in a small volume of starvation medium supplemented with 1.5 mCi of [35S]Cys·Met ProMix (Amersham) per ml. Cells were chased with chase medium (DME, 10% FCS, and 450 mg of cysteine and 300 mg of methionine per ml) and then disrupted in Triton lysis.

**Downloaded from http://jvi.asm.org/ on July 9, 2017 by guest**
buffer (50 mM Tris, pH 8.0; 150 mM NaCl; 0.1% sodium dodecyl sulfate [SDS]; 1 mM EDTA; 1% Triton X-100; protease inhibitor cocktail). Lysates were precleared with normal human immunoglobulin G (IgG; Sigma). Immunoprecipitation was accomplished using $2\text{mg}$ of anti-HIV IgIV, a human polyclonal antibody stock derived from HIV-1-seropositive individuals, and protein G-Sepharose beads (Pharmacia), followed by five washes with Triton lysis buffer.

RIPA of transfected COS-7 cells was performed in a similar manner. Approximately $5 \times 10^5$ transfected cells were suspended by treatment with phosphate-buffered saline (PBS)–EDTA (PBS, 0.5 mM EDTA, 5% dialyzed FCS) at 37°C for 15 min. Cells were then radiolabeled for 30 min in 50 $\mu$l of 1.5 mCi of $[35\text{S}]$Cys-Met ProMix per ml, washed once with Tris-saline (50 mM Tris [pH 8.0], 150 mM NaCl), and then lysed in RIPA buffer (50 mM Tris [pH 8.0], 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail). The lysates were precleared with 50 $\mu$g normal human immunoglobulin G and immunoprecipitated with anti-HIV IgIV. The resulting beads were washed with two washes of IP wash (50 mM Tris [pH 8.0], 150 mM NaCl, 0.1% NP-40), one wash of Tris-saline (50 mM Tris [pH 8.0], 150 mM NaCl), and one wash of 50 mM Tris (pH 6.8). We chose to assess the presence or absence of Pr55$^{gag}$ and p40 in transfected cells by RIPA rather than by Western blot because the HIV-IgIV antibody stock performs better in RIPA. However, similar results were also obtained by Western blotting (data not shown).

Detection of p40 in HIV-1-infected PBMC cultures was performed by pretreating $2 \times 10^6$ PHA-activated, CD8$^+$-depleted PBMC infected with vLAI, vATG, or vATC with or without indinavir for 2 h. Indinavir was provided by The Merck Company. Cells were starved for 30 min in starvation medium with or without indinavir and then radiolabeled with 200 $\mu$Ci of ProMix (with or without indinavir) for 1 h. Cells were washed and lysed using the same procedure as for COS-7 cells above, except that 50 $\mu$M indinavir was added to the lysis buffer and the lysate was immunoprecipitated with a polyclonal rabbit anti-p24 serum (NIAID ARRRP catalog no. 384) instead of anti-HIV IgIV to avoid possible detection of gp41.

In order to minimize background poliovirus bands, more stringent denaturing immunoprecipitation conditions were used to detect Pr55$^{gag}$ in transfected 293T cells infected with poliovirus. Transfected cells (see above) were infected with poliovirus by the addition of amplified poliovirus stock for the times shown in Fig. 1. Cells were then detached and radiolabeled in suspension for 45 min at 37°C. After radiolabeling, the cells were washed once in Tris-saline and then lysed by the addition of 35 $\mu$l of 2× SDS loading buffer (125 mM Tris, pH 6.8;...
The membrane was stained with Ponceau S (Bio-Rad), and the p40 blotted overnight at 15 V onto a polyvinylidene difluoride membrane (Bio-Rad). Nonprecipitated sample was electrophoresed by SDS-PAGE on a 12% gel and assayed by quantitation of total protein using a detergent-compatible protein assay kit (Bio-Rad) to fall within the 7- to 125-μg/ml range of the assay. Lysates were standardized by quantitation of total protein using a detergent-compatible protein assay kit (Bio-Rad) to fall within the 7- to 125-μg/ml range of the assay.

The predicted sizes the full-length transcript of each vector are given on the right.

20% glycerol; 2% β-mercaptoethanol; 4% SDS; 50 μg of bromophenol blue per ml. The lysate was incubated for 5 min at 100°C and then brought up in 1.5 ml of RIPA buffer. Samples were precleared with 50 μg of normal human immunoglobulin and then cleared with a mixture of 2 μg of anti-HIV IgV and 1 μg of polyclonal rabbit anti-CAT (Eppendorf 59, Inc.). The beads were then washed once with RIPA buffer, twice with EE wash (0.5 M NaCl; 50 mM Tris, pH 8.0; 5 mM EDTA; 5% sucrose; 1% NP-40), once with IP wash, once with Tris-saline, and once with 50 mM Tris (pH 6.8).

Immunoprecipitated samples were analyzed by comparison to The Benchmark Unstrained Protein Ladder (Life Technologies) by SDS-polyacrylamide gel electrophoresis (PAGE) on 8.5 or 12% gels.

**Northern blotting.** Cytoplasmic RNA was isolated (RNasey kit; Qiagen) from 293T cells transfected for 40 h with the plasmid vector indicated in Fig. 2. Each RNA (5 μg) was subjected to Northern blotting (NorthernMax-Gly Kit; Ambion) according to the kit instructions. Blots were probed with a biotinylated RNA probe generated by T7 transcription of vector pBhlg (Maxiscrip Kit; Ambion) with a 60:40 ratio of UTP to biotin-16-UTP (Boehringer). Detection was carried out using Ambion’s Biotrast Biodeetect Kit according to manufacturer’s instructions.

p24 ELISA. A commercially available anti-p24 ELISA kit (Coulter/Immunotech) was used to determine the levels of Gag in cell lysates and supernatants. Cells were washed twice with PBS-EDTA and then lysed in 300 μl of the lysis buffer provided with the kit. Cell lysates were diluted appropriately (typically, 1:100) to fall within the 7- to 125-μg/ml range of the assay. Lysates were standardized by quantitation of total protein using a detergent-compatible protein assay kit (Bio-Rad).

Edman microsequencing. Microsequencing of p40<sup>0⁰⁰⁰</sup> was performed by the Johns Hopkins University School of Medicine Biosynthesis and Sequencing Facility using a Perkin-Elmer/Applied Biosystems Procise 492 Protein Microsequencer. p40<sup>0⁰⁰⁰</sup> was isolated for microsequencing by immunoprecipitation of 5 × 10<sup>⁶</sup> vEB2-infected A42 EBV cells with HIV-IgV (see above). The immunoprecipitated sample was electrophoresed by SDS-PAGE on a 12% gel and blotted overnight at 15 V onto a polyvinylidene difluoride membrane (Bio-Rad). The membrane was stained with Ponceau S (Bio-Rad), and the p40<sup>0⁰⁰⁰</sup> band was excised and subjected to sequencing. Attempts to microsequence the 55-kDa species produced by vEB2-infected cells were performed in a similar manner except that the HIV-IgV antibody was covalently conjugated to protein G-Sepharose beads prior to immunoprecipitation to avoid elution of interfering immunoglobulin heavy chain.

**Virus culture.** For visualization of p40<sup>0⁰⁰⁰</sup> in HIV-1-infected cells, 2 × 10<sup>⁶</sup> activated/depleted PBMC were treated overnight with 1 ml of infectious supernatant collected from provirus-transfected COS-7 cells, incubated for 5 days, and then subjected to RIPA. For viral growth curves, provirus-transfected COS-7 supernatants were normalized by p24 ELISA so that roughly 50-ng equivalents (roughly 50 μl of supernatant) of p24 were added to 10<sup>⁶</sup> Jurkat cells or the stable cell lines Jurkat+sac or Jurkat+dac in a 1-ml volume in a 24-well plate. Cells were washed twice with 10 ml of culture medium 24 h postinfection. Supernatants were sampled periodically and subjected to p24 ELISA. Infected cultures were split as necessary.

Cesium-banded high-titer Mahoney type 1 poliovirus stock was the generous gift of Ellie Ehrenfeld (National Institutes of Health). In addition to being more easily transferrable than the HeLa cells typically used in poliovirus shutoff experiments, 293T cells were found to exhibit more-protracted shutoff kinetics than HeLa cells (data not shown). Amplified infectious poliovirus stocks were generated by adding ~0.1 μl of high-titer poliovirus stock (10<sup>11</sup> PFU/ml) to a confluent T-75 flask of 293T cells in 15 ml of culture medium for 24 h. The resulting cell lysate was clarified by centrifugation at 300 × g for 10 min, and 0.5 ml of the clarified lysate was used to infect confluent 4-cm<sup>2</sup> wells of transfected 293T cells.

**RESULTS**

Expression of Pr55<sup>0⁰⁰⁰</sup> in poliovirus-infected cells. To investigate the hypothesis that the HIV-1 gag ORF exhibits IRES activity, we undertook a series of experiments involving gag expression in poliovirus-infected cells. Because poliovirus mediates a rapid inhibition of cap-dependent translation in permissive host cells (23, 56), it has been widely used in the investigation of the cap dependency of translation of various mRNAs.

Figure 1 depicts a time course experiment in which 293T cells were first transfected with the indicated plasmid expression vector and then infected with poliovirus for various times. Bicistronic vector pCAT-GAG encodes the CAT reporter gene upstream of the Pr55<sup>0⁰⁰⁰</sup> gene, including a 77-bp segment of the native Pr55<sup>0⁰⁰⁰</sup> 5′ UTR. Control vector pEMCV-CAT encodes CAT under control of the EMCV IRES. After transfection and poliovirus infection, cells were radiolabeled with [<sup>35</sup>Si]Gys-Met for 45 min and subjected to immunoprecipitation with an anti-HIV–anti-CAT antibody mixture (Fig. 1A and B). Using this method, it was found that translation of EMCV IRES-driven CAT (Fig. 1A) and Pr55<sup>0⁰⁰⁰</sup> (Fig. 1B) were not significantly impacted by poliovirus-mediated shutoff of cap-dependent translation. In contrast, the cap-dependent translation of the upstream CAT ORF from vector pCAT-GAG declined substantially over the course of the infection. Bulk lysates of the transfected cells (Fig. 1C) displayed a pattern in which the translation of host proteins diminished with duration of poliovirus infection, while a characteristic pattern of poliovirus-specific bands began to appear.

A summary of PhosphorImager (Molecular Dynamics) quantitation of five independent poliovirus time course experiments is shown in Fig. 1D. In two of these experiments, cells were cotransfected with a CAT expression vector, pCAT, and a different bicistronic vector, pENV-ΔUTR-GAG (see below), instead of pCAT-GAG. The results suggest that the expression of Pr55<sup>0⁰⁰⁰</sup> from the constructs pCAT-GAG and pENV-ΔUTR-GAG was cap independent, a finding which supports the hypothesis that the HIV-1 gag gene encodes an IRES.

Northern blotting of pCAT-GAG. In addition to IRES activity, several possible explanations can be postulated to explain the expression of the second cistron of bicistronic constructs such as pCAT-GAG. These alternative scenarios include the existence of alternate splicing variants, RNA species derived from a cryptic promoter just upstream of the second cistron, or species of RNAs broken apart between the two cistrons. To rule out these possible explanations for the production of Pr55<sup>0⁰⁰⁰</sup> from construct pCAT-GAG, we performed Northern blots on cells transfected with pCAT-GAG (Fig. 2). This Northern blotting shows that pCAT-GAG produces only one transcript of each vector is given on the right.

![FIG. 2. Northern blotting of bicistronic vector pCAT-GAG. 293T cells were transfected with the plasmid expression vector indicated above each lane and subjected to Northern blotting. Kilobase RNA size markers are shown on the left. The predicted sizes the full-length transcript of each vector are given on the right.](http://jvi.asm.org/)

This Northern blotting shows that pCAT-GAG produces only one transcript of each vector.
major RNA which is the predicted size (3.4 kb) of the full-length bicistronic CAT-GAG mRNA. This result argues against the possibility that the production of Pr55\textsuperscript{ACE} is the result of a cryptic RNA species.

**Env-Gag bicistronic vectors.** To further investigate the IRES activity of the gag gene, we constructed a panel of bicistronic vectors. In these constructs, the HIV-1 env gene was used as an upstream reporter cistron, and the gag gene served both as the sequence to be tested for IRES activity and as a downstream reporter cistron. The env gene was chosen as an upstream reporter gene because it contains a large number of ATG codons and regions of highly structured RNA which would be expected to significantly disrupt scanning by any ribosomes bypassing the ATG codons within env (24). Expression of a novel 40-kDa gag isoform (p40), which results from translational initiation at an internal AUG codon, was also monitored in these experiments and will be discussed in greater detail below.

We first sought to determine what portion of the Pr55\textsuperscript{ACE} 5' UTR might be necessary for potentiating the translation of gag as a second cistron. Vector pENV-ΔUTR-GAG encodes only the env and gag reporter ORFs with the gag 5' UTR entirely omitted. This construct was found to express Pr55\textsuperscript{ACE} efficiently in poliovirus-infected cells (see above), suggesting that the native gag 5' UTR is dispensable for the IRES activity of the gag ORF. In order to determine what impact the native gag 5' UTR might have on gag's translation as a second cistron, we constructed vector pENV-UTR-GAG, which contains the entire Pr55\textsuperscript{ACE} 5' UTR encompassing most of the HIV-1 packaging (Ψ) signal. All three vectors were found to drive significant amounts of Pr55\textsuperscript{ACE} expression when compared to vector pGAG, which encodes only Pr55\textsuperscript{ACE} with an unobstructed vector-derived 5' UTR (Fig. 3A). This result confirms the observation that the native gag 5' UTR is not necessary to potentiate the expression of Pr55\textsuperscript{ACE} as a second cistron. Indeed, placement of the entire gag 5' UTR upstream of Pr55\textsuperscript{ACE} (in vector pENV-UTR-GAG) resulted in somewhat decreased levels of Pr55\textsuperscript{ACE} expression relative to the expression seen from vector pENV-ΔUTR-GAG. This decrease could have been due to activity of HIV-1 polyadenylation signal (found in the R domain of the HIV-1 LTRs), the presence of the primer binding site, or the presence of the viral dimerization initiation signal, since removal of these sequences (in pENV-Ψ-GAG) resulted in higher Pr55\textsuperscript{ACE} expression levels.

We generated several other vectors to investigate the conditions under which gag might function as a downstream cistron. In vector pΔENV-GAG, a frameshift was introduced near the 5' end of the env ORF of vector pENV-ΔUTR-GAG in order to demonstrate that translation of the env ORF is not required for translation of the downstream gag ORF. Vector pENVGAG encodes an in-frame fusion of the env ORF (lacking the native env stop codon) to the gag ORF. Expression of Pr55\textsuperscript{ACE} from this vector (Fig. 3A) argues against the possibility that the expression of Pr55\textsuperscript{ACE} from vector pENVGAG is due to translational reinitiation following env translation, since the env-gag and gag ORFs are cisternal. A control vector, pENVΔΔGAG, encodes the same Env-Gag fusion protein as vector pENVGAG except that the Pr55\textsuperscript{ACE} and p40 initiator codons have been mutated to ATC, a noninitiator Ile codon (79). As seen in Fig. 3A, mutation of these ATG codons abrogates Pr55\textsuperscript{ACE} and p40 production, suggesting that their expression is the result of translation initiation at these internal sites.

In a control vector, pENVCAT, all but the first 12 bp of the gag sequence of pENVGAG were replaced with the CAT reporter gene. As seen in Fig. 3B, pENVCAT does not drive expression of any CAT-sized species. This result suggests that the internal translational initiation at Pr55\textsuperscript{ACE} AUG\textsubscript{1} observed in vector pENVGAG is dependent on sequences within the gag ORF, since the CAT ORF was not sufficient to direct this activity.

In order to further exclude the possibility that Pr55\textsuperscript{ACE} and p40 expression from vector pENVGAG was the result of cryptic splicing or cryptic promoter usage, we generated a recombinant vaccinia virus vector, vENVGAG, which encodes the same Env-Gag fusion protein as plasmid pENVGAG. Because vaccinia viruses replicate in the cytoplasm of infected cells and
utilize virus-encoded transcription factors (59), it is unlikely that cryptic host promoters would have any activity in the context of a recombinant vaccinia virus. Splicing of vaccinia virus RNAs has not been reported. COS-7 cells infected with vvENVGAG exhibited Env-Gag, Pr55\(^{\text{gag}}\), and p40 bands (data not shown), indicating that expression of Pr55\(^{\text{gag}}\) and p40 from these vectors was not due to cryptic splicing or cryptic promoter usage.

These results confirm that sequences within the \textit{gag} ORF can promote internal initiation of translation of Pr55\(^{\text{gag}}\) even when the \textit{gag} ORF is placed downstream of large numbers of ATG codons and highly structured RNA sequences. This activity is a feature of sequences within the \textit{gag} ORF and functions even in the absence of the native \textit{gag} 5' UTR.

**Characterization of a 40-kDa \textit{gag} gene product.** A wide variety of plasmid and vaccinia virus expression vectors encoding Pr55\(^{\text{gag}}\) consistently display a minor \textit{gag} gene-specific 40-kDa band (Fig. 3 and 4 and data not shown). Because the \textit{gag} gene contains an isolated AUG at codon 142 that could initiate translation of a 40-kDa protein, we considered the possibility that p40 was the result of an internal initiation event at this AUG. To help determine its origins, we immunoprecipitated p40 from cells infected with vvB2, a vaccinia virus vector encoding Pr55\(^{\text{gag}}\), and subjected it to Edman microsequencing. This sequencing yielded the amino acid sequence NH\(_2\)-VHQAISPR\_A, a sequence identical to the predicted amino acid sequence of a polypeptide initiated at AUG\(_{142}\), assuming its initiator methionine was cleaved away (82; reviewed in reference 1). Pulse-chase experiments were also performed on cells infected with vvB2. Pr55\(^{\text{gag}}\) and p40 exhibited no precursor-product relationship, arguing against the concept that p40 was derived through proteolytic cleavage of Pr55\(^{\text{gag}}\) (data not shown).

We also examined a panel of plasmid expression vectors (Fig. 4) to verify that p40 was the result of an independent translation initiation event and not simply a by-product of Pr55\(^{\text{gag}}\). In two vectors, we independently mutated the Pr55\(^{\text{gag}}\) and p40 initiator ATG codons to ATC. We found that mutation of ATG\(_{142}\) to ATC abrogated p40 expression, suggesting that this codon is used as an initiator codon for p40. Although the mutation of Pr55\(^{\text{gag}}\) ATG\(_1\) to ATC prevented Pr55\(^{\text{gag}}\) expression, it did not abolish p40 expression. Similarly, introduction of a frameshift downstream of ATG\(_1\) did not abolish p40 expression. Thus, p40 can be expressed independently of Pr55\(^{\text{gag}}\) through translational initiation at gag codon 142.

**p40 expression in HIV-1-infected cells.** To determine whether internal initiation at AUG\(_{142}\) occurs during viral replication, we examined HIV-1-infected cells for the presence of p40. Answering this question was complicated by the activity of the HIV-1 protease. Met\(_{142}\) lies only six amino acids downstream of the eight-amino-acid protease recognition site (codons 129 to 136) at the matrix/capsid boundary. Thus, a single HIV-1 protease cleavage of Pr55\(^{\text{gag}}\) yields a protein nearly identical in size to p40. Similarly, a single cleavage of Pr55\(^{\text{gag}}\) at the capsid-p2 boundary results in a protein of approximately 40 kDa. For clarity, we will hereafter refer to the 40-kDa protein generated by initiation at AUG\(_{142}\) as p40\(_{\text{acc}}\) for “start at capsid.”

To determine whether p40\(_{\text{acc}}\) is produced in infected cells, we used indinavir (an HIV-1 protease inhibitor) to prevent the generation of Pr55\(^{\text{gag}}\) protease cleavage products (89). PHA-activated, CD8\(^+\)-depleted PBMC were infected with supernatants from COS-7 cells transfected with wild-type provirus pLAI, p40\(^{\text{acc}}\) knockout provirus pVATC (carrying an ATG\(_{142}\)→ATC [Met→Ile] mutation), or provirus pVATG, an isogenic wild-type control for pVATC (see Materials and Methods). Five days after infection, PBMC cultures were pretreated with 50 mM indinavir and then radiolabeled with \[^{35}\text{S}\]Cys-Met. Thus, all radiolabeled \textit{gag} gene products in the indinavir-treated cells were generated in the functional absence of HIV-1 protease. The labeled cells were then lysed and subjected to immunoprecipitation with anti-p24 serum.

Cells infected with wild-type viruses vLAI and vATG expressed 40-kDa products both in the presence and in the absence of indinavir (Fig. 5). The absence of a p24 band in cultures treated with indinavir suggests a strong inhibition of HIV-1 protease in these cultures. Although some 40-kDa Pr55\(^{\text{gag}}\) protease cleavage products accumulated in vATG-infected cells in the absence of indinavir, the ATG\(_{142}\)→ATC mutation in this virus completely ablated expression of p40\(_{\text{acc}}\), as shown by the absence of a 40-kDa band in the presence of indinavir. The results support the conclusion that HIV-1-infected cells produce small amounts of a novel 40-kDa protein, p40\(_{\text{acc}}\), through internal initiation at AUG\(_{142}\).
Rescue of impaired replication of a p40<sup>sac</sup>-deficient HIV-1 mutant. In order to investigate whether p40<sup>sac</sup> might play a role in viral replication in culture, we examined the replication kinetics of the p40<sup>sac</sup> knockout virus vATC. This mutant virus displayed dramatically impaired replication compared to wild-type vLAI or isogenic wild-type control vATG in Jurkat and CD8-depleted PBMC cultures (Fig. 6; also data not shown). In light of the fact that ATG<sub>142</sub> and the region surrounding it are highly conserved (60; HIV Sequence Database), even relatively conservative nonsilent mutations (such as the Met<sub>142</sub>→Ile mutation found in vATC) might be expected to be deleterious to viral growth through disruption of the normal functions of Pr<sub>55<sup> gag</sup></sub>.

To determine whether the delayed replication kinetics of vATC were due to its inability to produce p40<sup>sac</sup> or due to the amino acid substitution it carries in Pr<sub>55<sup> gag</sup></sub>, we generated a polyclonal stably transfected Jurkat line which expresses p40<sup>sac</sup> in trans. In the context of full-length HIV-1 mRNAs, translational initiation at Pr<sub>55<sup> gag</sup></sub> codon 142 would be expected to result in production of p40<sup>sac</sup> and a 145-kDa protein, p145<sup>sac-pol</sup>, by virtue of frameshifting at the HIV-1 "slippery" site. In an effort to provide appropriate levels of p40<sup>sac</sup> and p145<sup>sac-pol</sup>, we made a Rev- and Tat-inducible sac-pol vector, pSAC, based on the pLAI provirus. The vector expresses p40<sup>sac</sup> and p145<sup>sac-pol</sup> from the HIV-1 LTR. To block ribosome scanning and to monitor expression, we inserted the EGFP gene upstream of Pr<sub>55<sup> gag</sup></sub> codon 1, which was mutated from ATG to ATC to abrogate Pr<sub>55<sup> gag</sup></sub> expression. A fragment containing the HIV-1 RRE was placed downstream of the pol stop codon, as was a fragment carrying the CMV promoter, hygromycin B phosphotransferase, and a poly(A) signal. A Jurkat line stably carrying a control vector, p<sub>D</sub>SAC, in which the sac-pol gene was removed from pSAC, was also made. In the presence of Tat and Rev provided by an infecting virus, transcription and nuclear export of EGFP-p40<sup>sac</sup>-p145<sup>sac-pol</sup>mRNA expressed from these vectors should be induced. Consistent with this prediction, uninfected Jurkat cultures stably carrying these constructs displayed constitutive low levels of fluorescence (and low p24 ELISA signal for vector pSAC), while infected cultures of the lines showed a substantial population of highly fluorescent cells (data not shown). We named these cell lines Jurkat<sup>1</sup>sac and Jurkat<sub>D</sub>sac.

Although p40<sup>sac</sup> knockout virus vATC failed to grow to high titers in Jurkat and Jurkat<sub>D</sub>sac cultures, it was able to grow in the Jurkat<sup>1</sup>sac line (Fig. 6). When cultures were infected with much higher inoculums, vATC could eventually grow to high titers in Jurkat and Jurkat<sub>D</sub>sac cultures but with dramatically...
slower kinetics than seen in Jurkat+sac cultures (data not shown). Although vATC grew well in the Jurkat+sac line, growth was somewhat attenuated compared to wild-type virus vATG. This could reflect some impact of the Pr55<sup>gag</sup> ATG<sup>gag</sup>→ATC mutation in vATC, inappropriate timing of p40<sup>gag</sup> and p145<sup>sac-pol</sup> expression, or inappropriate levels of p40<sup>gag</sup> and p145<sup>sac-pol</sup> in the Jurkat+sac line.

These results are consistent with the concept that the defective growth phenotype of mutant virus vATC can be overcome, at least in part, by expression of p40<sup>gag</sup> and p145<sup>sac-pol</sup> in trans. Thus, expression of p40<sup>gag</sup> (or p145<sup>sac-pol</sup>) may be required for normal growth kinetics of HIV-1 in culture.

**DISCUSSION**

Our results demonstrate that the HIV-1 gag gene encodes an IRES. One line of evidence that Pr55<sup>gag</sup> can be translated in a cap-independent fashion via an IRES-like mechanism comes from experiments involving poliovirus-infected cells. Poliovirus mediates a rapid reduction in cap-dependent translation shortly after infection of permissive cells. This host cell translational shutoff coincides with the cleavage by poliovirus 2A protease of the eIF4G subunit of eIF4F (also known as the cap-binding complex), and the cleavage of poly(A) binding protein by poliovirus proteases 2A and 3C (44, 46). The IRES-mediated translation of poliovirus RNAs under these conditions is either unimpaired or enhanced. Poliovirus-infected cells (45) and extracts of cells infected with poliovirus have been widely used to demonstrate the presence of IRES activity within various mRNA sequences (23). As seen in Fig. 1, the translation of Pr55<sup>gag</sup> in cells transfected with bicistronic vector pCAT-GAG was insensitive to poliovirus-mediated shutdown. This supports the concept that Pr55<sup>gag</sup> can be translated through a cap-independent translation mechanism, such as an IRES, located within the gag ORF. Northern blots of cells transfected with pCAT-GAG were found to exhibit only one full-length mRNA species, arguing against the possibility that this cap-independent expression of Pr55<sup>gag</sup> was due to alternative splicing, cryptic promoter usage, or broken RNA species.

A separate line of evidence which suggests that gag encodes an IRES comes from a series of bicistronic vectors where cap-dependent ribosome scanning was obstructed by placement of the HIV-1 env ORF upstream of gag. As seen in Fig. 3A, placement of the env ORF upstream of gag does not abrogate Pr55<sup>gag</sup> expression. This result held true whether the native gag 5′ UTR was present or not, whether the env ORF was translated or not, and also when the env ORF was translated as an in-frame fusion with the gag ORF. Utilization of Pr55<sup>gag</sup> AUG<sub>1</sub> as an internal initiator codon required the presence of the gag ORF itself, as evidenced by the absence of expression of any CAT-length products by vector pENVCAT, in which CAT replaces most of the gag ORF downstream of env (Fig. 3B). These results argue against modified ribosome scanning mechanisms for Pr55<sup>gag</sup> or p40<sup>gag</sup> expression such as leaky scanning (49, 77), translational reinitiation (28, 54), ribosome shunting (72, 96), broken RNA, or ribosome “stutter” (18, 95). The expression of Pr55<sup>gag</sup> both from vaccinia virus vector vvENVGAG and plasmid vector pENVGAG argues against the possibility that expression of these proteins is due to cryptic promoter usage or cryptic splicing. Thus, two separate lines of evidence indicate that Pr55<sup>gag</sup> can be translated as an internal or downstream cistron by virtue of an IRES within the gag ORF.

In addition to potentiating the translation of Pr55<sup>gag</sup>, the HIV-1 gag IRES also facilitates translation of a novel 40-kDa protein, p40<sup>sec</sup>, whose translation initiates at an AUG codon near the 5′ end of the p24<sup>sac</sup> domain of Pr55<sup>gag</sup>. Several features of HIV-1 biology may have prevented the previous identification of this novel Gag isoform. As demonstrated in Fig. 5, intermediate products of the cleavage of Pr55<sup>gag</sup> by the HIV-1 protease are roughly 40 kDa in mass and can obscure the presence of p40<sup>sec</sup> (26). In past instances when infected cells treated with HIV-1 protease inhibitors have exhibited a 40-kDa band, the band may have been discounted as reflecting incomplete inhibition of protease activity by the drug (89, 92). The fact that p40<sup>sec</sup> is expressed at much lower levels than Pr55<sup>gag</sup> has also perhaps hidden it from identification even when Pr55<sup>gag</sup> has been expressed in the absence of HIV-1 protease (25, 57, 73).

Despite its low abundance, our results suggest that p40<sup>sec</sup> may play a vital role in viral replication in vitro. HIV-1 mutant virus vATC, which fails to generate p40<sup>sec</sup>, and, presumably, p145<sup>sac-pol</sup> replicates with dramatically impaired growth kinetics in cultured cells (Fig. 6). The replication defect of vATC can be partially overcome by the expression of p40<sup>sec</sup> or p145<sup>sac-pol</sup> in trans, suggesting that the defect stems, at least in part, from the mutant’s inability to generate these novel proteins. This observation is consistent with the concept that p40<sup>sec</sup> and/or p145<sup>sac-pol</sup> play an important role in viral replication in culture.

It has been shown that sequences within the native gag 5′ UTR can inhibit cap-dependent translation of reporter proteins. The Tat-responsive element (TAR) is an RNA secondary structural element that is responsible for binding HIV-1 Tat protein and may play a role in the inhibition of cellular interferon-induced double-stranded RNA-activated kinase (12, 22). The TAR is located at the 5′ end of all HIV-1 RNA transcripts. Because the presence of secondary structure at or near the 5′ end of RNAs reduces the accessibility of the 5′ cap to eIF4F (49), it is thought that this feature of HIV-1 mRNAs can inhibit their cap-dependent translation (64, 78). Other stable HIV-1 RNA structural elements involved in regulating the splicing, polyadenylation, dimerization, packaging, primer binding, and reverse transcription of viral RNAs are thought to inhibit translation by blocking the scanning of 40S ribosomes (and their associated cofactors) through the gag 5′ UTR (29, 57). Thus, a possible function of the HIV-1 gag IRES might be to serve as a mechanism to bypass the structural barriers to cap-dependent translation by recruiting ribosomes directly to the gag ORFs.

It has been demonstrated that infection with HIV-1 can arrest cells in the G<sub>2</sub> phase of the cell cycle (reviewed in reference 13). The fact that cap-dependent translation is inhibited during some phases of the cell cycle (8) led Vagner and colleagues to speculate that an HIV-1 IRES might allow for enhanced HIV-1 gene expression under growth-arrest conditions where the cap-dependent translation of host cell mRNAs is inhibited (91). Recent reports have found that the IRESs of hepatitis C virus (HCV) and host mRNAs encoding ornithine decarboxylase and p58<sup>ITSLRE</sup> are significantly more active
during the G₂/M phase of the cell cycle than in the G₁ or S phases (11, 34, 69). The possible interplay between HIV-1 cell cycle arrest and IRES utilization deserves future investigation.

In addition to the IRESs found in the 5' UTRs of a variety of plus-stranded RNA viruses (3, 9, 31, 42, 43, 66, 88) and a number of eukaryotic cellular mRNAs (6, 27, 37, 38, 61, 62, 84, 90), IRESs have recently been described in several type C retroviruses. These include MMLV (91), Friend murine leukemia virus (FMLV) (4, 19), the Rat VL30 region of the Harvey murine sarcoma virus (5), and avian reticuloendotheliosis virus type A (53). A variety of other retroid family members, including human T-cell leukemia virus type 1 (2), human foamy virus (51), LINE-1 and VL30 (two groups of endogenous vertebrate retro-elements) (39, 52, 56, 87), and Drosophila R2 retrotransposable elements (30), have also been shown to exhibit possible IRES-like activities. These data, when taken together with our observation in this report that HIV-1 encodes an IRES, suggest the possibility that IRES-mediated translation is a typical feature of the retroviridae.

Murine leukemia viruses encode two different Gag proteins. Glyco-Gag, a nonstructural transmembrane glycoprotein important for viral infectivity, is initiated from a CUG codon located upstream of and in frame with the AUG-initiated gag gene (67, 68). The IRES of MMLV has been shown to be situated entirely within the viral glyco-gag ORF in a 126-bp sequence just upstream of the gag AUG. The MMLV IRES was shown to initiate translation only at the gag AUG, whereas the FMLV IRES, which has not been fully localized, was shown to be capable of directing translation of both Gag and Glyco-Gag.

Localization of the MMLV gag IRES to a region entirely contained within an ORF was striking in that it (and perhaps the IRES of FMLV) was the first naturally occurring (10) IRES entirely contained within a translated ORF. More recently, a similar ORF-localized IRES has been described for the host mRNA encoding p10<sup>PITSLRE</sup> and p58<sup>PITSLRE</sup> (11). The HIV-1 Pr<sup>55<sub> gag</sub></sup> IRES activity described here is thus conceptually similar to the MMLV and PITSLRE IRESs in that it is located entirely within an ORF and is similar to the FMLV IRES in that it can direct translation of two different gag gene products.

Miele and colleagues have argued that the HIV-1 packaging signal does not encode an IRES (57). Using a series of bicistronic constructs in which the entire gag 5' UTR and as much as the first 200 nt of the gag ORF were inserted between two reporter cistrons, these investigators found no evidence for potentiation of translation of the downstream reporter cistron. In contrast to Miele and colleagues' findings, a recent report from Ohlmann and colleagues has found that sequences in the 5' UTR of the gag gene of another primate lentivirus, SIVmac, do exhibit IRES activity (63). Thus, it may be that IRES activity is typical of the lentiviridae, while the exact location of sequences required for the activity varies.

Perhaps the most surprising feature of the HIV-1 gag IRES is that it directs translational initiation at (or, conceivably, beyond) its 5' border. Precedents for translational enhancement by sequence elements located 3' of the ORF they control include the mRNAs of certain viruses (14, 85, 94), and the 3' UTR of the mRNA for the β subunit of mitochondrial H<sup>+</sup>-ATP synthase (40). Important future work on this topic will include identification of the minimal sequence elements required for HIV-1 gag ORF IRES activity and the potential mechanisms underlying its surprising, seemingly backward, arrangement.

Das et al. have demonstrated that constitutive expression of a small inhibitor RNA (IRNA) molecule isolated from yeast (15) can effectively protect cultured cells from productive poliovirus infection (17) without significantly impacting cell growth or morphology. The suppression of poliovirus replication in cells expressing this IRNA is likely the result of competition between the IRNA and the poliovirus IRES for access to host proteins required for viral IRES-mediated translation (16, 93). It is thus formally possible that similar antagonism of the HIV-1 gag IRES could disrupt the replication of HIV-1 or other IRES-bearing retroviruses.

ACKNOWLEDGMENTS

We thank Mark J. Selby, Lucy Carruth, Tim Tobery, Monika Hermannova, and Xiao Fang Yu for valuable advice. We also thank Elizabeth K. Flynn for field trials of the Benchmark Protein Ladder which facilitated the discovery of p<sup>55<sub> gag</sub></sup>. We are especially grateful to Ellie Ehrenfeld and Richard Lloyd for patiently providing advice and materials which were crucial for the completion of this research. This work was supported by NIH grants AI28108 and AI37924 and by a grant from the Merck Trust.

REFERENCES


eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature 334:320–325.