JOURNAL OF VIROLOGY, Apr. 1992, p. 1915–1923
0022-538X/92/041915-09$02.00/0

Vol. 66, No. 4

Expression of Intracisternal A-Particle-Related Retroviral Element-Encoded Envelope Proteins Detected in Cell Lines

FRANK U. REUSS†

Zentrum für Molekulare Biologie, Universität Heidelberg, Im Neuenheimer Feld 282, D-6900 Heidelberg, Germany

Received 1 July 1991/Accepted 18 December 1991

Intracisternal A-particle (IAP) retrotransposons of rodents express *gag* and *pol* proteins for assembly of intracellular viruslike particles but lack an *env* gene. The recently described IAP-related family of retroviral elements contains a reading frame with close resemblance to retroviral *env* genes (IAPEs) (F. U. Reuss and H. C. Schaller, J. Virol. 65:5702–5709, 1991). I now report the analysis of cellular IAPE mRNAs and detection of IAPE *env* proteins. IAPE elements are transcribed in cell lines NH15-CA2 and AtT20. Four major transcripts of 4.2, 3.9, 2.8, and 1.3 kb are detected and characterized by probes specific for defined regions of the cloned IAPE-1 cDNA. The 2.8-kb mRNA is shown to lack *gag* and *pol* genes but comprises an *env* gene and U3 region, as expected for a subgenomic *env* mRNA. Polymerase chain reaction amplification and cloning of such mRNAs confirmed the absence of *gag* and *pol* genes 5′ from the *env* gene and implicates *env* mRNA generation by a splicing event. A polyclonal anti-IAPE *env* antiserum, raised against a bacterial IAPE-*env* fusion protein, specifically detects N-glycosylated *env* proteins of 91 kDa or less in cell lines positive for IAPE mRNA. IAPE *env* proteins of different sizes represent independent translation products. After inhibition of N-glycosylation, *env* proteins in the size predicted from the *env* gene sequence or smaller are present. These results provide evidence that putative IAPE *env* proteins are synthesized in vivo. Envelope protein expression by an IAP-related retroviral element identifies IAPEs as a possible missing link between IAP retrotransposons and retroviruses.

The infectivity of retroviruses is critically dependent on envelope proteins present on the surface of the viral particle. These virus-encoded *env* proteins are responsible for recognition of the appropriate host cell by interaction with a cellular receptor (20) and for entrance of the virus into the cell by fusion with the cell membrane (34).

The three main protein components of the retroviral particle are encoded by the proviral DNA in the order 5′-*gag-pol-env*-3′. Whereas *gag* and *pol* proteins are translated from a full-length transcript of the provirus, expression of the *env* protein is dependent on a splicing event that removes *gag* and *pol* genes from the primary transcript and generates a subgenomic *env* mRNA. Splicing either moves an initiating methionine codon at the beginning of the *env* reading frame close to the 5′ end of the mRNA (18) or fuses the strong initiation codon plus several codons of the *gag* gene in-frame with *env* gene sequences (10). The efficiency of this splicing process is balanced by suboptimal splicing signals (14) to guarantee the presence of full-length genomic RNAs and expression of *gag* and *pol* genes. When *env* mRNA is translated on membrane-bound polyribosomes, the *env* protein is inserted into the membrane of the endoplasmic reticulum.

The envelope protein becomes N-glycosylated during synthesis and assembles into an oligomeric structure. During transport to the cell surface, *env* proteins are endoproteolytically cleaved by a cellular protease (33). The two resulting polypeptides, the surface attachment subunit and the membrane-anchoring subunit, remain connected to each other by disulfide bonds. Upon arrival at the surface of the host cell, *env* proteins recognize intracellular viral capsids and interact with them. This results in budding and release of viral particles from the membrane of the host cell. Free particles undergo a maturation process and are subsequently able to infect new target cells.

The intracisternal A-particle (IAP) group of defective retroviruses lack the *env* gene (see reference 16 for review). Large numbers of IAP proviral elements present in rodent genomes direct the synthesis and assembly of intracellular viruslike particles. Due to the lack of a functional envelope gene in IAP elements and possibly other defects as well, no extracellular infectious virions are produced. Instead, the assembling particles bud into the endoplasmic reticulum of the host cell.

The recently described IAP envelope-coding (IAPE) class of retroviral elements and cDNAs from *Mus musculus* (26) are related to known IAP genomes but contain an additional gene which lies in a position characteristic of retroviral *env* genes. Also, the length and structural features of the deduced protein are characteristic of retroviral envelope proteins. It is, however, not known whether the encoded IAPE *env* protein is expressed in vivo.

The aim of this study was to determine whether the IAPE *env* gene is used to direct the synthesis of protein. For this purpose, several mouse cell lines were tested for transcription of IAPE mRNA and the structure of transcripts was analyzed by using region-specific IAPE probes. A 2.8-kb mRNA exhibited the structure of a subgenomic *env* mRNA. Radioimmunoprecipitation of cellular proteins with anti-IAPE *env* serum specifically detected N-glycosylated IAPE *env* proteins of the expected sizes.

MATERIALS AND METHODS

Cell lines. I used the pituitary cell line AtT20 (4), NIH 3T3 fibroblasts, P19 embryonal carcinoma cells (19), and the (mouse neuroblastoma × rat glioma) × mouse neuroblas-

† Present address: Department of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111.

toma hybrid-hybrid cell line NH15-CA2 (12), a gift of B. Hamprecht. NH15-CA2[a] and NH15-CA2[b] are morphologically distinguishable sublines of the originally aneuploid NH15-CA2 line that were passaged independently. NH15-CA2[a] cells from passage 17 were not passaged more than 10 times before use. NH15-CA2[b] cells, in contrast, were continuously propagated. NIH 3T3 and NH15-CA2 cells were cultured in Dulbecco's modified Eagle's medium (DMEM), and AtT20 and P19 cells were cultured in RPMI 1640 medium (Boehringer, Mannheim, Germany) supplemented with 10% fetal calf serum (FCS).

IAPE-specific probes. All IAPE probes consisted of RNA complementary to the sense strand of cDNA IAPE-1 (26) and were synthesized with RNA polymerase T7 or SP6 (Promega) in an in vitro transcription reaction. The names of the probes and nucleotide positions of IAPE-1 covered are GAG (1 to 418), POL (1055 to 1270), ENV1 (2185 to 2441), ENV2 (2916 to 3113), LTR1 (3650 to 3851), and LTR2 (3846 to 4021). Control probe pGEM2PL represents the polylinker sequence of the pGEM2 plasmid (nucleotides 10 to 86) present in IAPE probes.

Northern (RNA) blotting and hybridization. RNA was isolated from cell lines by acid guanidinium thiocyanate extraction as described before (3). The poly(A)-containing RNA fraction [poly(A)-RNA] was enriched by affinity chromatography on oligo(dT)-cellulose columns. Five micrograms of poly(A)-RNA was electrophoresed per lane on agarose-formaldehyde gels and transferred to Hybond-N membranes (Amersham). As size markers run on the same gel, a set of IAPE-1-derived sense transcripts in the size range of 0.6 to 5.2 kb were produced. These transcripts, complementary to the antisense probes used, were generated by in vitro transcription from plasmid pGEM2-IAPE-1 (nucleotides 1 to 3851) and several of its derivatives after enzymatic linearization at different defined sites. Hybridization was performed under high-stringency conditions at 60 or 65°C in 40% formamide–5× SSC (1× SSC is 0.15 M sodium chloride plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate (SDS)–1× Denhardt's solution (0.2 mg of Ficoll, 0.2 mg of polyvinylpyrrolidone, and 0.2 mg of bovine serum albumin per ml)–1 mM EDTA–5 mM sodium phosphate buffer (pH 6.8)–0.2 mg of yeast tRNA per ml–0.2 mg of salmon sperm DNA per ml for 16 h and washed three times for 30 min at 65°C in a buffer consisting of 0.1× SSC–0.1% SDS. The amounts of mRNA loaded in each lane were controlled by hybridization to a ^{32}P-labeled actin probe. After autoradiography, the probe was removed by incubating the membrane in 5 mM Tris hydrochloride (pH 8)–2 mM EDTA–0.1× Denhardt's solution at 90°C. Successful removal of the probe was tested by autoradiography.

PCR. Starting with 2 μg of NH15-CA2 poly(A)-RNA, IAPE transcripts were specifically reverse transcribed by 20 U of avian myeloblastosis virus enzyme (Promega) with primer P1 (5′-GAAAGTCTAGAGAATAGCCC-3′) complementary to positions 1756 to 1775 of IAPE-1. The products were subjected to a 25-cycle polymerase chain reaction (PCR) (27) with primers P1 and P2 (5′-GAGAGAATTCCT GAGAAGCAAGGAA-3′, positions 3991 to 4008 of IAPE-1) at an annealing temperature of 57°C, extension temperature of 72°C, and denaturation temperature of 94°C for 1 min each in TRB buffer (50 mM KCl, 10 mM Tris hydrochloride [pH 8.3], 2.5 mM $MgCl_2$, 0.1% gelatin). Next, 2.5 U of *Thermus aquaticus* DNA polymerase (Perkin Elmer) was added per 100-μl sample. Fragments identified by Southern blot analysis of the separated products with ^{32}P-labeled oligonucleotide P3 (5′-GCAGGGAATTCATCAAGGATGGTG-3′, positions 1644 to 1667 of IAPE-1) as the probe were isolated from the gel, and recessed 3′ ends were filled in with Klenow enzyme. These fragments were cleaved with *Eco*RI and *Xho*I, cloned into pGEM2, and sequenced.

DNA sequencing. DNA sequencing was performed on plasmid DNA by the dideoxy chain termination method (29), with modified T7 DNA polymerase (Sequenase; USB) and T7 or SP6 primers. Both strands of a fragment were sequenced completely.

Computer-assisted sequence analysis. Sequence analysis and comparisons were performed with the University of Wisconsin Genetics Computer Group sequence analysis software package (5). The program MULTALIGN (17) was used to compare multiple sequences.

Expression of recombinant IAPE *env* proteins and production of antisera. IAPE-0 cDNA (nucleotides 1929 to 2669 of IAPE-1) plus added *Eco*RI linker (26) was cloned into the *Eco*RI site of vector pEX34b (9). This plasmid expresses foreign antigens as fusion proteins with the N-terminal 99 amino acids of the MS2 bacteriophage RNA polymerase. Protein synthesis upon heat induction at 42°C in bacterial strain K-12DH1ΔtrpEA (1) was tested. Preparative isolation of the MS2-*env* fusion protein was carried out via enrichment of bacterial inclusion bodies and solubilization in 8 M urea as described before (30). Successive dilution of the urea concentration to 4 and 0.6 M and subsequent centrifugation (Beckman L8-M, SW28; 20°C, 30 min, $10^5 \times g$) was used to fractionate the proteins according to differential solubility. The precipitate from the 0.6 M urea solution was suspended in TBS (25 mM Tris hydrochloride [pH 7.5], 150 mM NaCl), and about 0.1 mg of fusion protein was injected subcutaneously into rabbits as an emulsion with Freund's adjuvant. Booster injections were given at 4-week intervals, and blood was taken 7 days after the boost. Sera were tested for recognition of IAPE *env* protein by Western immunoblotting experiments with fusion protein β-galactosidase (βgal)-*env* or control βgal produced from plasmid pEX2-IAPE-0 or pEX2 (32) as the antigen. βgal-*env* and βgal proteins were isolated in parallel as described before (30). Before further use, the antiserum was purified by using two affinity columns. The first column was prepared by coupling a bacterial lysate containing the enriched MS2 portion of the recombinant protein to CNBr-activated Sepharose CL-4B (Pharmacia). The flowthrough from this procedure was then absorbed either to a βgal-*env* fusion protein-coupled column (for control serum) or to a βgal protein-coupled column (for anti-*env* serum). The flowthrough from these columns differed only in the absence (control serum) or presence (anti-*env* serum) of anti-IAPE *env* antibodies, as was shown in Western blot experiments. A comparison of control and anti-*env* sera served as a criterion for IAPE *env*-specific recognition.

^{35}S labeling and immunoprecipitation. Cells were grown in six-well plates (35-mm-diameter well; Falcon) in complete medium including 10% FCS to about 70% confluence, washed with methionine-free DMEM (GIBCO) plus 2 mM glutamine, incubated twice with the same medium for 15 min each, and labeled for 1 h with 0.4 mCi of [^{35}S]methionine (>1,000 Ci/mmol; Amersham) per ml. In the case of pulse labeling, incubation was for 15 min at 37°C with 0.6 mCi of [^{35}S]methionine per ml and followed by a chase in DMEM supplemented with 10% FCS for different time periods. Cells were then washed twice with ice-cold 0.1 M Tris hydrochloride (pH 6.8), scraped from the plates, lysed in sample buffer (2% SDS, 62.5 mM Tris hydrochloride [pH 6.8], 5% β-mercaptoethanol, 5% glycerol, 0.02% bromophenol blue), soni-

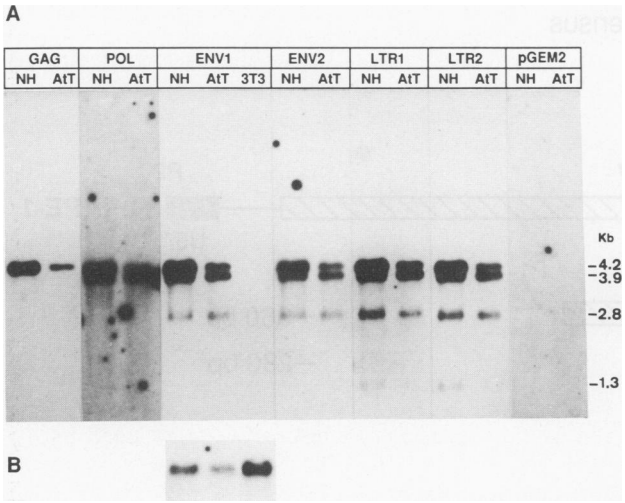

FIG. 1. Analysis of cellular IAPE transcripts. Poly(A)-RNA isolated from cell lines NH15-CA2[b] (NH), AtT20 (AtT), and NIH 3T3 (3T3) was separated in an agarose-formaldehyde gel and transferred to a Hybond-N membrane. This membrane was successively hybridized to (A) probes GAG, POL, ENV1, ENV2, LTR1, LTR2, and pGEM2PL (pGEM2) and (B) to a chick actin probe. The size of the RNAs is indicated.

cated, and heated at 95°C for 3 min. After dilution with 9 volumes of RIPA buffer (150 mM NaCl, 10 mM sodium phosphate [pH 7.5], 1% Nonidet P-40, 0.5% deoxycholate), immunoprecipitations were performed at 4°C for 7 to 10 h, first with the control serum and then with the anti-*env* serum. Equivalents of 2 μl of undiluted serum were bound to protein A-Sepharose CL-4B (Pharmacia) in both cases. Extensive washing in RIPA–0.1% SDS and final denaturation of the pellet in sample buffer (3 min, 95°C) preceded SDS-polyacrylamide gel electrophoresis (PAGE) and autoradiography. Tunicamycin incubations were done for 3 h with concentrations of 25 (AtT20), 10 (3T3), or 5 μg (NH15-CA2) of tunicamycin per ml before and during radioactive labeling.

RESULTS

Analysis of IAPE transcripts. We recently described the sequence of novel IAPEs isolated from the neural rat-mouse hybrid cell line NH15-CA2 (26). Retroviral elements highly related to the isolated cDNAs are present in the mouse genome but are not found in the rat genome (26). To analyze the cellular transcripts of these IAPE elements, poly(A)-RNA was isolated from several mouse-derived cell lines. I used the cell line NH15-CA2[b], adenopituitary-derived AtT20 cells, NIH 3T3 fibroblasts, and P19 embryonal carcinoma cells. The RNAs were separated on a formaldehyde-agarose gel and transferred onto a Hybond-N membrane. This membrane was consecutively hybridized to different radiolabeled probes derived from the cloned cDNA IAPE-1 (26). The probes were specific for the 5′ untranslated region and *gag* gene (GAG), *pol* gene (POL), *env* gene (ENV1 and ENV2), and U3/R region (LTR1 and LTR2) of IAPE-1 and consisted of short 100 to 400-nucleotide RNAs complementary to the transcripts.

Among the four cell lines tested, only NH15-CA2 and AtT20 were found to contain IAPE-specific transcripts (Fig. 1). IAPE RNA could not be detected by this method in NIH 3T3 and P19 cells with any of the probes used. Hybridization of NIH 3T3 RNA with probe ENV1 is shown here as an example. This result proves that IAPE elements are transcribed only in some mouse cell lines and suggests a similar regulation mechanism by *cis*-acting signals localized in the long terminal repeat (LTR) of IAPEs, as has been shown for other retroviral elements (35).

Both NH15-CA2 and AtT20 cells were found to contain four major IAPE mRNA species of different sizes. This is consistent with the existence of several classes of IAPE elements in the mouse genome (26). The transcripts detected were of the same sizes in both cell lines. Depending on the IAPE probe used for hybridization, different mRNAs were detected (Fig. 1).

The largest IAPE mRNA is 4.2 kb in size. Two lines of evidence identify this mRNA species with the original template of the cloned cDNA IAPE-1 (26). First, the 4.2-kb mRNA was the only RNA recognized by all IAPE-1-derived probes. Second, as judged by size, the 4.2-kb mRNA was the only RNA species from which the 4,021-nucleotide cDNA IAPE-1 could have been derived.

The 3.9-kb mRNA is characterized by the lack of sequences complementary to the GAG probe. This suggests the presence of a deletion in the *gag* and *pol* genes larger than the one found in cDNA IAPE-1.

A 2.8-kb mRNA is visualized by the ENV and LTR probes only and must therefore be considered as having a role in *env* gene expression. The 1.3-kb transcript is detectable with LTR probes exclusively. Detection of the 1.3-kb mRNA in AtT20 was possible after prolonged exposure of the autoradiograph (data not shown). No major IAPE transcript of 7 to 8 kb, characteristic of complete retroviral RNAs, could be detected in the cells analyzed.

Splicing deletes *gag* and *pol* genes from *env* mRNA. All retroviral *env* proteins are translated from a spliced subgenomic mRNA devoid of *gag* and *pol* genes (25). The only IAPE mRNA in consideration for this purpose is the 2.8-kb species, which was shown to lack *gag* and *pol* genes. In order to determine the exact nature of the deletion, I investigated the structure of such RNAs in the region 5′ to the *env* coding region. An IAPE-specific cDNA was synthesized from NH15-CA2 poly(A)-RNA with oligonucleotide primer P1 from the *env* gene of IAPE-1 (Fig. 2A). To gain access to the 5′ end of IAPE cDNAs, I took advantage of the fact that retroviral mRNAs carry identical repeats (R) at both ends. A second oligonucleotide primer, P2, was chosen from the predicted repeat region at the 3′ end of IAPE-1. The IAPE-specific cDNA was used as the template for a PCR with P1 and P2 as the primers. The products were separated on an agarose gel and probed with a second *env*-specific oligonucleotide, P3, in a Southern blot experiment. One main 360-bp product and a minor product of 280 bp were detected (Fig. 2A). The major fragment was cloned, and five of the clones, SP1 to SP5, were randomly chosen for sequence analysis.

The nucleotide sequences of all five SP clones were highly related, but no two of the clones were exactly identical (Fig. 2B). The degree of divergence from a theoretical consensus sequence was determined for each clone and ranges from 2 to 16 nucleotide exchanges. A comparison of these sequences with cDNA IAPE-1 reveals that in all SP clones, a specific loss of sequence information for *gag* and *pol* genes has occurred at the same position (Fig. 2A). I analyzed the sequences in cDNA IAPE-1 surrounding the deletion points and found that the DNA sequence in the vicinity of both sites is in agreement with known consensus sequences for rodent

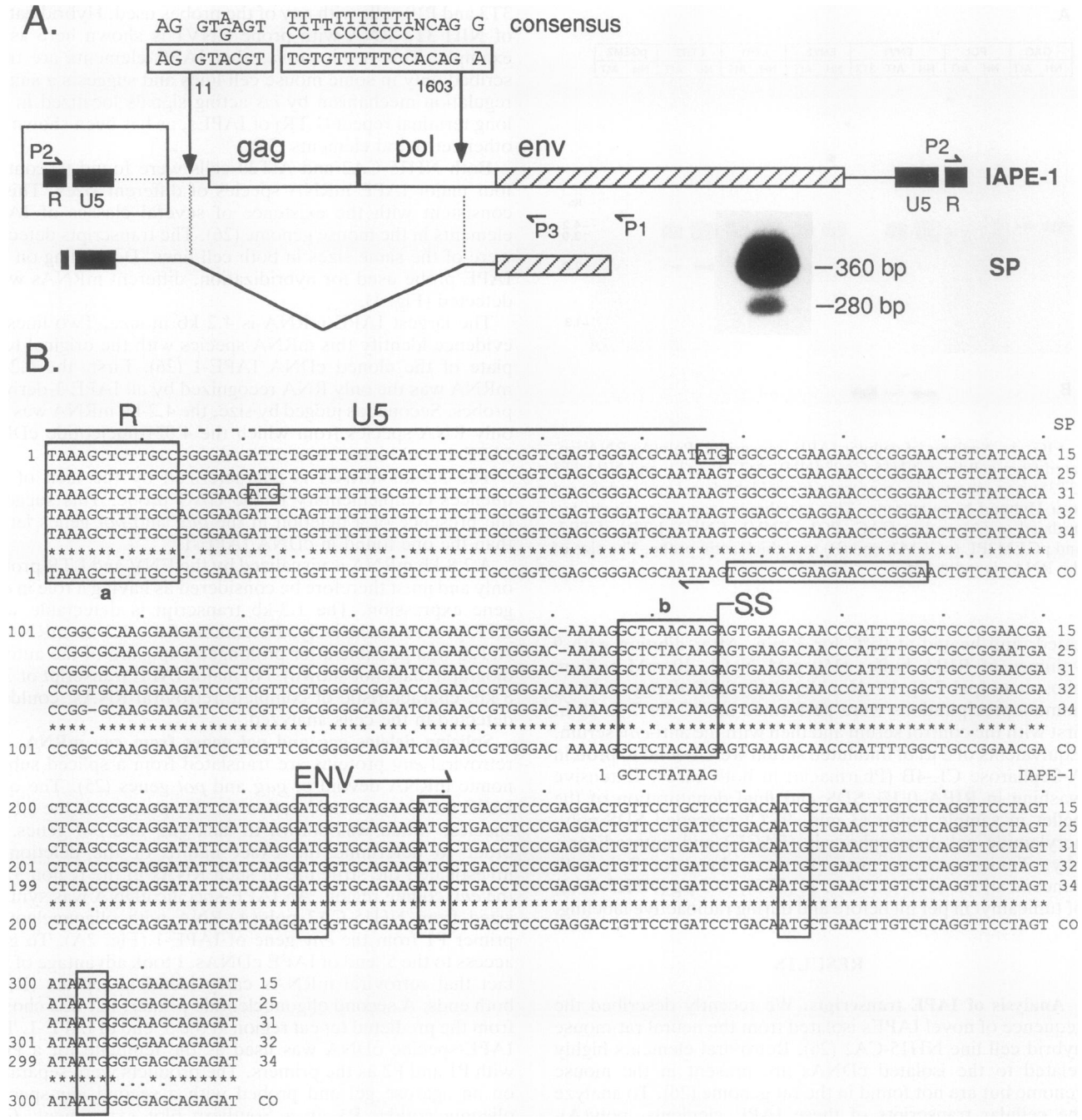

FIG. 2. (A) Enzymatic amplification of IAPE mRNAs deleted for *gag* and *pol* genes. NH15-CA2[b] poly(A)-RNA was reverse-transcribed with primer P1 and enzymatically amplified with primers P1 and P2. Products were identified by Southern blot analysis with radiolabeled probe P3 (see autoradiograph). The structural organization of the product (SP) is schematically drawn in comparison with cDNA IAPE-1 (26). The missing 5′ region of IAPE-1 is boxed. IAPE-1 sequences in the vicinity of the deletion points are presented above the schematic drawing and compared with rodent splice consensus sequences (31). (B) Nucleotide sequence comparison of PCR-amplified subgenomic IAPE cDNAs. Nucleotide sequences of five randomly chosen clones (SP 15, 25, 31, 32, and 34) are compared and a consensus sequence (CO) is derived. Sequences similar to the IAPE-1 repeat region (a, nucleotides 4009 to 4021) and the 5′-terminal region of IAPE-1 (b, nucleotides 1 to 10) are boxed. The splice site (SS) is marked. A box in the consensus sequence indicates the region complementary to the 3′ end of mouse tRNA^{Phe}. The first four ATG codons in the *env* gene are labeled. R and U5 mark the presumed repeat and U5 regions, respectively.

splice donor and acceptor sites (31). This suggests that posttranscriptional deletion of *gag* and *pol* genes from primary IAPE transcripts occurred via a splicing mechanism. The apparent deletion in SP clones comprised nucleotides 11 to 1603 of IAPE-1 and placed the *env* gene in proximity to the 5′ end of the mRNA.

The novel 5′ exon was identified as part of an IAPE mRNA by the fact that it carries the presumed R region of

```
A.
tRNA^Phe 3`-ACCACGGCTTT--GGGCCCT-5`
          ::: ::::::::  :::::::
IAPE     5`-TGGCGCCGAAAGACCCGGGA-3`

B. MIA14  TGGTGACCGAAT--TCCGGGA
          *** * *****    ******
   IAPE   TGGCG-CCGAAAGACCCGGGA
          *** * ******  *******
   H18    TGGTG-CCGAAA--CCCGGGA
```

FIG. 3. tRNA^{Phe} and IAP primer-binding sites. (A) Complementarity of mouse tRNA^{Phe} (11) to the putative primer-binding site in IAPE; (B) nucleic acid comparison between putative primer-binding sites of mouse IAP element MIA14 (21), mouse IAPE (SP clones), and Syrian hamster IAP-like element H18 (24).

IAPE-1 (nucleotides 4009 to 4021) at its 5′ end and nucleotides 1 to 10 of IAPE-1 at its 3′ end (Fig. 2B). The sequence of IAPE-1 in the corresponding regions is not exactly identical to that of the SP clones. In addition to variations confined to one SP clone only, all SP clones differ from IAPE-1 at the position corresponding to nucleotide 7 of IAPE-1. This means that none of the mRNAs represented by SP clones was generated by splicing from the IAPE-1 mRNA.

In two of the five SP clones, depending on the sequence of the 5′ end of the mRNA, the *env* gene is preceded by novel methionine codons in the same reading frame. Although a possible involvement of these codons in *env* protein synthesis cannot be ruled out at present, the variable position of these methionine codons and their absence from most SP clones render this unlikely. In contrast, the methionine codons at positions 1, 5, 17, and 27 of the putative *env* protein are conserved in all SP clones and hence seem to be of more importance for its expression or function. In summary, these data provide evidence for the existence of a subgenomic IAPE *env* mRNA that is probably generated by splicing.

tRNA^{Phe} as a putative IAPE replication primer. The SP clones described above not only represent part of a subgenomic IAPE mRNA, but in addition extend the known IAPE-1 cDNA sequence in the 5′ direction. In the 5′ exon, a previously unidentified part of IAPEs, approximately 130 nucleotides in size, is found (Fig. 2B). This region contains a 20-bp sequence with strong complementarity to the 3′ end of mouse tRNA^{Phe} (Fig. 3). DNA stretches complementary to this tRNA species, thought to bind tRNA^{Phe} as the primer for minus-strand DNA synthesis during replication, have been found exclusively in IAPs (21). The identity of the putative tRNA primers in IAPEs and IAPs emphasizes the close relationship between both groups of retroviral elements.

Immunoprecipitation of IAPE *env* proteins. To investigate the expression of IAPE *env* proteins in cell lines containing IAPE mRNA, I raised a polyclonal antiserum against amino acids 90 to 336 of the putative *env* protein (26), expressed as a bacterial fusion protein. The antiserum was purified by removal of antibodies specific for bacterial proteins and for the portion of the fusion protein unrelated to IAPE *env*. This antiserum was used in a two-step procedure to immunoprecipitate proteins from total-cell lysates of cell lines metabolically labeled with [^{35}S]methionine. A control immunoprecipitation with the serum, specifically cleared of anti-*env* antibodies by absorption to an affinity column (control serum), was used to define and reduce the unspecific background. The precleared lysate was then immunoprecipitated with anti-*env* serum in a second step (Fig. 4).

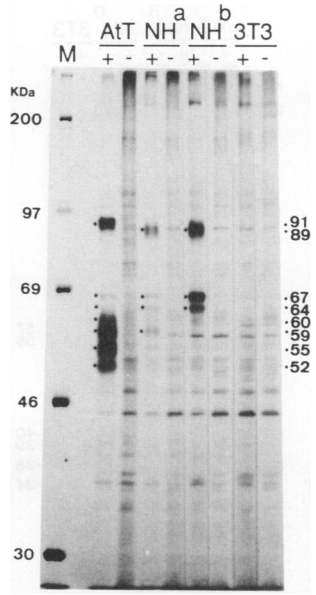

FIG. 4. Immunoprecipitation of IAPE *env* proteins from cell lines. Cell lines were metabolically labeled with [^{35}S]methionine for 1 h. Whole-cell lysates of cell lines AtT20 (AtT), NH15-CA2^a (NH^a), NH15-CA2^b (NH^b), and NIH 3T3 (3T3) were immunoprecipitated first with control serum (−) and then with anti-IAPE *env* serum (+). Proteins specifically precipitated by the anti-*env* serum are marked (◆). The molecular mass of these proteins and marker proteins (lane M) is indicated.

In the IAPE-transcribing cell lines AtT20 and NH15-CA2, but not in the IAPE mRNA-negative NIH 3T3 cells, a set of proteins were specifically recognized by anti-*env* antibodies (Fig. 4). The exact *env* protein pattern differed between cell lines AtT20 and NH15-CA2 and even between different sublines of the originally aneuploid NH15-CA2 line but was composed of characteristic combinations of eight proteins. AtT20 and NH15-CA2 cells could be distinguished by the size of the largest *env* protein, 91 and 89 kDa, respectively. Proteins of lower molecular mass (67, 64, 62, 59, 55, and 52 kDa) were all represented in AtT20. Of these, only the 67- and 64-kDa species were found in NH15-CA2^b, whereas 67-, 64-, and 59-kDa species are present in NH15-CA2^a. The expression of *env* proteins was significantly lower in NH15-CA2^a than in NH15-CA2^b, as judged by the relative intensity of unspecifically precipitated proteins. In summary, these results demonstrate that IAPE *env* genes are used for the expression of several protein species in cell lines.

IAPE *env* proteins are N-glycosylated. Retroviral *env* proteins generally become N-glycosylated during translocation into the endoplasmic reticulum (6, 13). The IAPE *env* gene also predicts a protein product with 16 potential N-glycosylation sites (26). In order to analyze whether the proteins identified above are modified by N-glycosylation and to investigate the role of differential N-glycosylation in the generation of different IAPE *env* proteins, the cells were cultured in the presence of tunicamycin before and during ^{35}S labeling (Fig. 5). As a consequence of N-glycosylation inhibition, the molecular mass of all IAPE *env* proteins decreased. A 56-kDa protein was common to all cell lines and is the largest *env* protein expressed in AtT20 and NH15-CA2^a cells. NH15-CA2^b cells also contained an additional larger protein of 57 kDa. These values are in accordance with the calculated molecular mass of 62.0 kDa for the

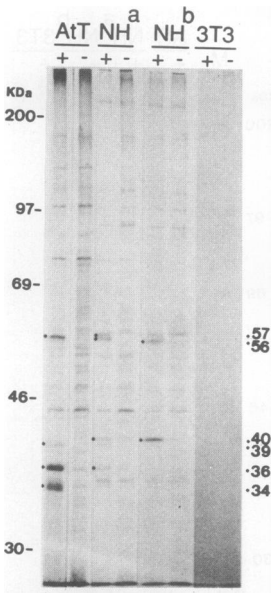

FIG. 5. Immunoprecipitation of IAPE *env* proteins after tunicamycin treatment. Cell lines AtT20 (AtT), NH15-CA2[a] (NH[a]), NH15-CA2[b] (NH[b]), and NIH 3T3 (3T3) were preincubated with tunicamycin 3 h before and during [^{35}S]methionine labeling in order to inhibit N-glycosylation. Further treatment was as described in the legend to Fig. 4.

IAPE *env* precursor protein after cleavage of the signal peptide (residues 1 to 25). Additional smaller proteins of 39.5, 36, and 35 kDa in AtT20, 40 and 36 kDa in NH15-CA2[a], and 40 kDa in NH15-CA2[b] cells were also present. These results provide evidence that all detectable IAPE *env* proteins are modified by N-linked glycosylation. N-glycosylation seems to contribute to the heterogeneity of IAPE *env* proteins. The apparent number of IAPE *env* protein species could be reduced by inhibition of N-glycosylation, e.g., in AtT20 from seven to four, indicating that different numbers or qualities of N-linked oligosaccharides are added to protein chains of identical length.

IAPE *env* proteins are independent translation products. IAPE *env* proteins of different sizes are detectable in various cell lines. The question of whether these proteins are independent translation products or derive from a precursor protein by proteolytic cleavage was addressed by a pulse-chase experiment in AtT20 cells. A 15-min [^{35}S]methionine pulse was followed by a chase period of variable length in complete medium including FCS.

The results (Fig. 6A) showed that even if the cells were lysed immediately after a 15-min pulse, all IAPE *env* proteins previously identified in AtT20 cells were already present. The chase procedure, during the time periods investigated, did not identify any precursor-product relationships between the *env* proteins. The individual IAPE *env* proteins, however, were subject to size reduction and clearance from the cell. The apparent molecular mass of all proteins was slightly decreased during the chase period, indicating intracellular processing. In parallel to this, the intensity of individual protein bands decreased with time after synthesis. The quantity of each of the major proteins was measured by densitometric analysis of the autoradiograph (Fig. 6B). In general, large proteins were detectable in the cells for longer periods than smaller proteins. The largest protein, gp91, was visible even after a 4-h chase, and the smallest protein, gp52, was hardly detectable after a 1-h chase.

I conclude from these data that the different N-glycosylated IAPE *env* proteins in AtT20 cells are independent

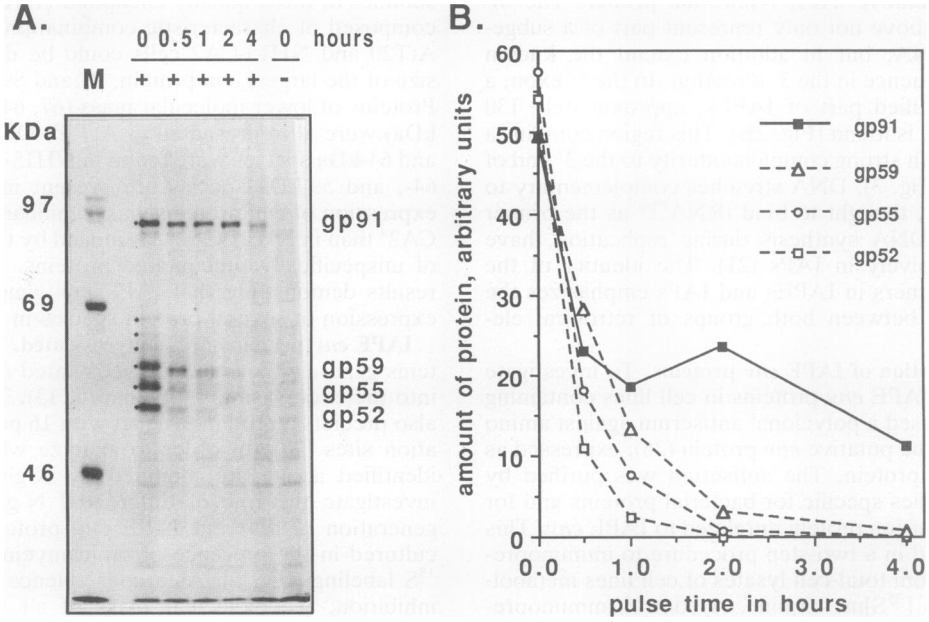

FIG. 6. Pulse-chase analysis of IAPE *env* proteins in AtT20 cells. (A) Cells were pulse-labeled for 15 min and chased in complete medium for 0, 0.5, 1, 2, or 4 h. Immunoprecipitation was done from whole-cell lysates with control serum (−) or anti-IAPE *env* serum (+). IAPE *env* proteins are marked (◆); the apparent molecular mass of *env* proteins and marker (lane M) proteins is indicated. (B) The autoradiograph in panel A was scanned densitometrically, and the peak area was calculated. The values for gp91, gp59, gp55, and gp52 are plotted against chase time.

translation products and undergo intracellular processing. gp91 is distinguished from all other IAPE *env* proteins by its size and prolonged intracellular presence.

DISCUSSION

The recently characterized IAPE class of mouse retroviral elements are related to known IAP elements but, in contrast to these, code for an envelopelike protein (26). In this study, I followed the expression of IAPE *env* genes from the characterization of different IAPE mRNAs and identification of a subgenomic *env* mRNA to the detection of different IAPE *env* proteins in mouse cell lines.

IAPE mRNAs were detected in the neural cell line NH15-CA2, from which IAPE cDNAs were originally isolated, and in cell line AtT20, derived from the adenopituitary gland. Whereas IAPs and other retroviruses have been identified by electron microscopy in AtT20 cells (36), no similar studies have been done with NH15-CA2 cells. The analysis of mRNA from both cell lines detected four different IAPE transcripts. This observation is analogous to that from studies on IAP elements, which express several different mRNAs in a particular cell line (15). Three of the IAPE mRNAs, 4.2, 3.9, and 2.8 kb in size, contain the *env* gene and thus should be considered for their potential role in *env* protein expression.

The 4.2-kb mRNA, as suggested by its size and hybridization pattern, gave rise to the cloned 4,021-bp cDNA IAPE-1 (26). It is characterized by a large truncation of *gag* and *pol* genes, reducing the length of both genes from 5 kb in the full-length IAP element MIA14 (21) to 1.45 kb. The 3.9-kb mRNA has an even larger truncation of the *gag* and probably also of the *pol* gene. As a consequence, no functional *gag* and *pol* proteins can be expressed from either mRNA.

Both the 4.2- and 3.9-kb mRNAs are presumably transcribed from genomic IAPE elements of corresponding sizes and hybridization patterns, previously identified as 4-kb and 3.7-kb *Eco*RI fragments, respectively (26). Based on the existence of genomic IAPE elements in the range of 7.5 kb that represent potential full-length IAPE elements, mRNAs of similar sizes were expected. Such transcripts, however, could not be identified in any of the cells analyzed. As a direct consequence of these findings, IAPE *gag* and *pol* proteins, generally translated from such full-length mRNAs, and the viral particles generally assembled from these constituents can at best be expected in small amounts. This finding does not, however, exclude potential transcription of full-length IAPE mRNA in other cell lines. It has been shown for mouse IAPs that certain cell lines transcribe truncated IAP elements exclusively, whereas large amounts of full-length IAP mRNA are present in others (15).

The synthesis of *env* proteins from the 4.2- and 3.9-kb mRNA species is rendered unlikely by the presence of 1.1- to 1.4-kb parts of the *gag* and *pol* genes 5′ to the *env* gene. Retroviral *env* proteins are generally translated from a specifically spliced subgenomic mRNA devoid of *gag* and *pol* genes (25). The 2.8-kb mRNA is suggested to be a spliced subgenomic *env* mRNA by the following results. First, differential hybridization revealed that is lacks *gag* and *pol* genes, but the *env* gene and U3/R region are present. Second, the exact position and nature of the *gag* and *pol* deletion was defined, and the nucleotide sequences in the vicinity of the deletion sites are in agreement with known consensus splice donor and splice acceptor sites. Third, no class of IAPE elements in the mouse genome has been detected that is similar to the 2.8-kb mRNA in size and hybridization pattern and at the same time has a copy number sufficient to generate a minimum of five different mRNAs (26). Taken together, these arguments suggest that a splicing mechanism removes the *gag* and *pol* genes from the primary IAPE transcripts.

The exact structure of the 1.3-kb IAPE mRNA, detectable by LTR probes only, has not yet been analyzed in detail. Cells infected with mouse mammary tumor virus, a virus known to have certain *env* protein features in common with IAPEs, contain a spliced 1.7-kb mRNA. A presumed regulatory protein, encoded 3′ to the *env* gene, can be translated from this mRNA (37). It is tempting to speculate that the 1.3-kb IAPE mRNA may serve to express two small overlapping open reading frames encoded in a 270-bp region between the *env* gene and the 3′ LTR of IAPEs (26).

One spliced subgenomic *env* mRNA species of defined length was identified. However, this mRNA species is known to be a complex mixture of mRNAs with heterogeneous sequences and hence is the basis for the expression of a large family of *env* proteins. By using a polyclonal antiserum against part of the postulated surface attachment subunit, IAPE *env* proteins in eight different size classes were identified. Translation in the absence of N-glycosylation produces proteins of higher electrophoretic mobility. The largest of these proteins, present in all cells, has an apparent molecular mass of 56 kDa and is interpreted as the full-size envelope precursor protein. The measured size is in sufficient agreement with the value of 62 kDa calculated for the *env* protein with the signal peptide cleaved. The divergence between both values may arise by abnormal behavior in SDS-PAGE, as known for the structurally related *env* protein of mouse mammary tumor virus. The calculated molecular mass of 66.2 kDa for the unglycosylated mouse mammary tumor virus *env* precursor also differs from the measured 60-kDa apparent molecular mass in SDS-PAGE (6, 22). The difference amounts to 6 kDa in both cases.

The smaller *env* proteins are thought to arise as a consequence of premature stop codons in the *env* reading frame of mRNA subpopulations. As cloned IAPE cDNAs confirm, IAPE mRNAs differ from each other at about 2% of the nucleotide positions (26). Some of these mutations result in amino acid exchanges in the deduced protein or even in the creation of stop codons, leading to chain termination (26). The alternative explanation, generation of these proteins from a larger precursor molecule by proteolytic processing, is unlikely, considering the results of pulse-chase experiments in AtT20 cells. Not all of the smaller proteins are of the same size in different cell lines. A differential activation of particular IAPE subsets or independent accumulation of mutations in these cell lines may be responsible for this phenomenon.

As predicted by the IAPE *env* cDNA sequence (26) and common to all retroviral *env* gene products, N-glycosylation increases the apparent molecular mass of IAPE *env* proteins. After N-glycosylation, the apparent molecular mass of the largest protein species, the putative *env* protein precursors, shifts to 91 kDa in AtT20 and 89 kDa in NH15-CA2 cells. These glycoproteins most likely consist of the 56- or 57-kDa protein that is modified with oligosaccharides at most of the 16 potential N-glycosylation sites. The size difference between gp91 and gp89 may well be explained by cell line-specific glycosylation, as shown for other viral glycoproteins (8, 23).

The putative envelope protein precursor gp91 is further distinguished from the smaller IAPE *env* proteins by its relative stability between 0.5 and 2 h after synthesis. It is

after that time that a decrease in the amount of gp91 can be noticed. The nature of this step may be the same as in other retroviral *env* protein precursors that are generally cleaved into two subunits by a cellular protease during transport to the cell surface (6, 33). With IAPE gp91, however, no *env* protein cleavage product becomes visible during the chase periods tested. One possible explanation is a combination of inefficient cleavage and degradation, as in human immunodeficiency virus type 1 (38).

Synthesis of *env* proteins distinguishes IAPEs from the retrotransposons expressing only *gag* and *pol* proteins (2). Since the production of *env* proteins is confined to functional retroviruses or vertically transmitted proviruses highly related to them (28), IAPEs may represent the missing link between IAPs and retroviruses.

Retroviral *env* proteins are a main constituent of extracellular viral particles and a prerequisite for infectivity of the virus. Thus, if there is full-length IAPE RNA as well as IAPE *gag* and *pol* proteins, IAPE *env* proteins may allow the assembly of an as yet unknown infectious IAPE virus. The relationship of IAPEs to IAPs suggests the additional possibility that an IAP-type viral capsid could interact with IAPE *env* proteins to assemble an IAPE pseudotyped IAP retrovirus. Testing for IAPE and IAP components in cell culture supernatants and animals is expected to provide further information about a potential IAPE retrovirus.

ACKNOWLEDGMENTS

This work was carried out in the laboratory of H. Chica Schaller (Zentrum für Molekulare Biologie, Heidelberg, Germany). I thank U. Klingmüller, H. Schaller, and R. M. Flügel for helpful discussions, N. Schek for advice on immunoprecipitations, K. K. Lueders for correspondence, and L. Runkel for reading the manuscript. Oligonucleotides were synthesized by W. Reiser and A. Helfer, and expression plasmids were kindly provided by E. Beck and P. Schloss.

This work was supported by DFG (SFB 317).

REFERENCES

1. **Bernard, H.-U., E. Remaut, M. V. Hershfield, H. K. Das, and D. R. Helinski.** 1979. Construction of plasmid cloning vehicles that promote gene expression from the bacteriophage lambda pL promoter. Gene **5**:59–76.
2. **Boeke, J. D., and V. D. Corces.** 1989. Transcription and reverse transcription of retrotransposons. Annu. Rev. Microbiol. **42**: 403–434.
3. **Chomczynski, P., and N. Sacchi.** 1987. Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
4. **Cohen, A. I., J. Furth, and R. F. Buffet.** 1975. Histologic and physiologic characteristics of hormone secretory transplantable adrenal tumors in mice and rats. Am. J. Pathol. **33**:631–651.
5. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
6. **Dickson, C., and M. Atterwill.** 1980. Structure and processing of the mouse mammary tumor virus glycoprotein precursor $Pr73^{env}$. J. Virol. **35**:349–361.
7. **Dickson, C., J. A. Puma, and S. Nandi.** 1976. Identification of a precursor protein to the major glycoproteins of mouse mammary tumor virus. J. Virol. **17**:275–282.
8. **Etchinson, J., and J. Holland.** 1974. Carbohydrate composition of the membrane glycoprotein of vesicular stomatitis virus grown in four mammalian cell lines. Proc. Natl. Acad. Sci. USA **71**:4011–4014.
9. **Felleisen, R., M. Q. Klinkert, and E. Beck.** 1988. Schistosoma mansoni: localisation of antigenic regions on the 31 kilodalton diagnostic protein. Mol. Biochem. Parasitol. **30**:19–26.
10. **Ficht, T. A., L.-J. Chang, and C. M. Stolzfus.** 1984. Avian sarcoma virus *gag* and *env* gene structural protein precursors contain a common amino-terminal sequence. Proc. Natl. Acad. Sci. USA **81**:362–366.
11. **Gauss, D. H., and M. Sprintzl.** 1983. Compilation of tRNA sequences. Nucleic Acids Res. **11**:r1–r52.
12. **Heumann, R., M. Öcalan, V. Kachel, and B. Hamprecht.** 1979. Clonal hybrid cell lines expressing cholinergic and adrenergic properties. Proc. Natl. Acad. Sci. USA **76**:4674–4677.
13. **Hunter, E., and R. Swanstrom.** 1990. Retrovirus envelope proteins. Curr. Top. Microbiol. Immunol. **157**:187–253.
14. **Katz, R. A., and A. M. Skalka.** 1990. Control of retroviral RNA splicing through maintenance of suboptimal processing signals. Mol. Cell. Biol. **10**:696–704.
15. **Kuff, E. L., and J. W. Fewell.** 1985. Intracisternal A-particle gene expression in normal mouse thymus tissue: gene products and strain variability. Mol. Cell. Biol. **5**:474–483.
16. **Kuff, E. L., and K. K. Lueders.** 1988. The intracisternal A-particle gene family: structure and functional aspects. Adv. Cancer Res. **51**:183–276.
17. **Krueger, M., and G. Osterburg.** 1983. On the alignment of two or more molecular sequences. Comp. Prog. Biomed. **16**:61–70.
18. **Majors, J. E., and H. E. Varmus.** 1983. Nucleotide sequencing of an apparent proviral copy of *env* mRNA defines determinants of expression of the mouse mammary tumor virus *env* gene. J. Virol. **47**:495–504.
19. **McBurney, M. W., and B. J. Rodgers.** 1982. Isolation of male embryonal carcinoma cells and their chromosome replication patterns. Dev. Biol. **89**:503–508.
20. **McDougal, J. S., M. S. Kennedy, J. M. Sligh, S. P. Cort, A. Mawle, and J. D. A. Nicholson.** 1986. Binding of HTLV-III/LAV to $T4^+$ T-cells by a complex of a 110 K viral protein and the T4 molecule. Science **331**:382–385.
21. **Mietz, J. A., Z. Grossman, K. K. Lueders, and E. L. Kuff.** 1987. Nucleotide sequence of a complete mouse intracisternal A-particle genome: relationship to known aspects of particle assembly and function. J. Virol. **61**:3020–3029.
22. **Moore, R., M. Dixon, R. Smith, G. Peters, and C. Dickson.** 1987. Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of *gag* and *pol*. J. Virol. **61**:480–490.
23. **Nakamura, K., and R. Compans.** 1979. Host cell- and virus strain-dependent differences in oligosaccharides of hemagglutinin glycoproteins of influenza A viruses. Virology **95**:8–23.
24. **Ono, M., H. Toh, T. Miyata, and T. Awaya.** 1985. Nucleotide sequence of the Syrian hamster intracisternal A-particle gene: close evolutionary relationship of type A particle gene to types B and D oncovirus genes. J. Virol. **55**:387–394.
25. **Pawson, T., R. Harvey, and A. E. Smith.** 1977. The size of Rous sarcoma virus mRNAs active in cell-free translation. Nature (London) **268**:416–420.
26. **Reuss, F. U., and H. C. Schaller.** 1991. cDNA sequence and genomic characterization of intracisternal A-particle-related retroviral elements containing an envelope gene. J. Virol. **65**:5702–5709.
27. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
28. **Salmons, B., G. Knedlitschek, N. Kennedy, B. Groner, and H. Ponta.** 1986. The endogenous mouse mammary tumour virus locus Mtv-8 contains a defective envelope gene. Virus Res. **4**:377–389.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
30. **Schloss, P., I. Hermans-Borgmeyer, H. Betz, and E. D. Gundelfinger.** 1988. Neuronal acetylcholine receptors in Drosophila: the ARD protein is component of a high-affinity α-bungarotoxin binding complex. EMBO J. **7**:2889–2894.
31. **Shapiro, M. B., and P. Senapathy.** 1987. RNA splice junctions of different classes of eucaryotes: sequence statistics and functional implication in gene expression. Nucleic Acids Res. **15**: 7155–7174.

32. **Stanley, K. K., and J. P. Luzio.** 1984. Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. EMBO J. **3:**1429–1434.
33. **Stein, B. S., and E. G. Engleman.** 1990. Intracellular processing of the gp160 HIV-1 envelope precursor. J. Biol. Chem. **265:** 2640–2649.
34. **Stein, B. S., S. D. Gowda, J. D. Lifson, R. C. Penhallow, K. G. Bensch, and E. G. Engleman.** 1987. pH-independent HIV entry into CD4-positive T-cells via virus envelope fusion to the plasma membrane. Cell **49:**659–668.
35. **Stoltzfus, C. M.** 1988. Synthesis and processing of avian sarcoma retrovirus RNA. Adv. Virus Res. **35:**1–38.
36. **Tooze, J., S. Tooze, H. Haisma, and J. Hilgers.** 1985. AtT20 pituitary tumour cells contain mouse mammary tumour virus and intracisternal A-type particles in addition to murine leukemia virus. Eur. J. Cell Biol. **39:**224–231.
37. **van Ooyen, A. J. J., R. J. A. M. Michalides, and R. Nusse.** 1983. Structural analysis of a 1.7-kilobase mouse mammary tumor virus-specific RNA. J. Virol. **46:**362–370.
38. **Willey, R. L., J. S. Bonifacio, B. J. Potts, M. A. Martin, and R. D. Klausner.** 1988. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. Proc. Natl. Acad. Sci. USA **85:**9580–9584.