Expression of dihydrofolate reductase, and of the adjacent EIb region, in an Ad5-dihydrofolate reductase recombinant virus

Kathleen L.Berkner and Phillip A.Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 1 November 1983; Revised and Accepted 23 January 1984

ABSTRACT

A gene with the Ad2 MLP and first leader, and appropriate RNA processing signals (splicing, polyadenylation) positioned around a mouse DHFR cDNA clone was substituted for the EIa region of Ad5, and virus stocks of Ad5 (DHFR-I) were prepared on 293 cells. A DHFR RNA of the expected size and structure was expressed late after infection of 293 cells by Ad5 (DHFR-I), at levels comparable to that of other Ad5 late messages. Although this DHFR mRNA was translated as efficiently as other Ad late mRNAs in vitro, it was only poorly translated in vivo.

The substitution of the DHFR gene for the Ad5 EIa region results in aberrant transcriptional activity in the adjacent EIb sequences. The transcriptional levels of the EIb 1kb message were down approximately 10-fold. In addition, a novel pIX-encoding mRNA was produced, generated by the splicing of the Ad first late leader onto sequences 14 bp upstream from the pIX initiation codon. This new mRNA was found to be potently efficient for translation both in vivo and in vitro.

INTRODUCTION

The adenoviral system provides distinct advantages for adaptation as vectors for the introduction of genes of interest into mammalian cells. The large genome size (30-35 kb), coupled with Ad-transformed cell lines that complement deletion mutant Ad DNAs (1, 2), allow for insertions of large fragments of foreign DNA. DNA substituted into the EI or the EIII region (3) can be propagated as homogenous virus stocks. During the late stage of infection, the level of transcription from the major late promoter (MLP) is high and most of the cellular protein synthesis (4) is inhibited. Thus, late in Ad-infection viral proteins account for 90% of the new protein synthesis. A foreign gene placed under Ad late control might therefore be expressed at very high levels.

Recently, a panel of genes encoding dihydrofolate reductase (DHFR), which are capable of transforming chinese hamster ovary (CHO) cells deficient in DHFR ($DHFR^-$) to a $DHFR^+$ phenotype, were described (5, 6). These genes are composed of modular units: a segment containing the Ad2 MLP positioned

© IRL Press Limited, Oxford, England.

upstream for initiation of transcription, a segment containing a cDNA clone of mouse DHFR mRNA, and segments with appropriate splicing and polyadenylation signals. In CHO-lines transformed by the modular DHFR gene, methotrexate (MTX) selection resulted in an amplification in gene copy of almost 10^3-fold. These cells synthesized DHFR mRNAs with the anticipated structures (6).

In these studies, an Ad5 recombinant with a modular DHFR gene substituted for the EIa region was constructed. This particular construct contained the MLP positioned upstream from the DHFR segment. Virus stocks of this recombinant were prepared by propagation on the EI-complementing 293 cell line (1), and expression of DHFR in infected cells was analyzed. While the levels of DHFR RNA synthesis were comparable to other late Ad viral RNAs, protein synthesis of DHFR was not correspondingly high. During these studies, we observed unexpected alterations in the expression of the EIb unit, which was adjacent to the DHFR gene in this recombinant.

MATERIALS AND METHODS

Cell culture and viral infections

Cell lines were grown in Dulbecco's modified Eagle's medium (DME, Microbiological Associates) and 10% fetal calf serum (293 cells) or in 10% calf serum (JW34 HeLa). DUKX-B1, a CHO DHFR$^-$ line (from L. Chasin, 7), was grown in α^- media (Microbiological Associates) supplemented with 10 mg/l each of adenosine, deoxyadenosine, and thymidine, and 10% fetal calf serum.

Preparation of virus stocks of Ad5 (DHFR-I)

The DHFR modular gene chosen for insertion into Ad5 DNA, pAdD26SV(A), was prepared by Randal Kaufman, and its structure has been described in detail elsewhere (5). Its insertion into the Ad5 genome is illustrated in Figure 1. The indicated DNA fragments were purified by electrophoresis through agarose and isolation by treatment with NaI and glass beads (9). XbaI-cleaved pDHFRI DNA (2 ug) was ligated with 1.3 ug of a sucrose-gradient purified fragment of XbaI-cleaved 309 DNA (10), i.e. XbaI-A (3.8-100 mu) and then ethanol precipitated with 10 ug salmon sperm DNA and NaCl (to 0.25 M). The DNA was transfected onto 293 cells using calcium phosphate precipitation (11), as previously described (3), and plaques were isolated after 10 days.

RNA analyses

Total polyA+ RNA was prepared as described by Seed and Goldberg (12). Northern analysis of these RNAs was as previously described (3).

For S1-nuclease analysis, a probe for DHFR-I RNA was prepared by cleaving

pAd5 (DHFR-I) with AccI endonuclease and 5' end-labeling with [γ-^{32}P]ATP and T4 polynucleotide kinase. The DNA was then digested with Eco RI endonuclease, and the 1.2 kb Eco RI-AccI fragment was gel purified. A second probe was prepared from SmaI endonuclease digested viral Ad5 (DHFR-I)DNA, 5' end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The left most SmaI fragment (4.5 kb) was purified by agarose gel electrophoresis. DNA/RNA hybridizations were performed by incubating labeled probe (10 ng) with 50 ng polyA+ RNA at 50°C or 52°C for 7 hr. The samples were then digested with S1 nuclease ($4x10^3$ u) for 30' at 45°C (13), and the hybrids were phenol extracted and ethanol precipitated.

Protein analysis of infected cells

Early (6 hr) or late (14-24 hr) after viral infection (at mois of 20), cells were labeled with ^{35}S-methionine, as previously described (3). Samples were electrophoresed through Laemmli SDS-containing gels (14) with 15% acrylamide and 0.086% N,N'-methylene bisacrylamide, for 3 hr at 130 V.

Immunoprecipitation was with Staphylococcus aureas as the immunoadsorbent (15). Cellular extracts were precleared for 16 hr at 4°C with washed Staph A (at 2% w/v) in 2 uM phenyl methyl sulfonyl floride, and then incubated with antisera for 5 hr at 4°C. Staph A was added (10-serum volumes), incubated at 4°C for an additional 4 hr, and then centrifuged. The pellet was washed twice in RIPA buffer containing 0.5 M NaCl, then once in RIPA with 0.15 M NaCl. The antisera used here were a monoclonal, α-protein IX, kindly provided by Connie Cepko and described elsewhere (16), and α-DHFR, provided by Bruce Dolnick and Joseph Bertino.

Reagents

All restriction enzymes were purchased from New England Biolabs or from Bethesda Research Laboratories. T4 polynucleotide kinase was from Collaborative Research, T4 polynucleotide ligase from New England Biolabs, and S1 was from Miles. [^{35}S]-met and [α-^{32}P]dNTPs were purchased from New England Nuclear, and [γ-^{32}P]ATP was from ICN.

RESULTS

Preparation of the Ad5 (DHFR-I) recombinant

An Ad5 recombinant, Ad5 (DHFR-I), with most of the EIa (1-3.8 mu) region substituted by a DHFR gene, was constructed according to the protocol detailed in Fig. 1. This plasmid contains the leftmost 350 bp of Ad5 [sequences that are required for Ad replication and packaging (17, 18)], the Ad2 MLP and entire first leader with the 5' splice site, and a 3' splice site

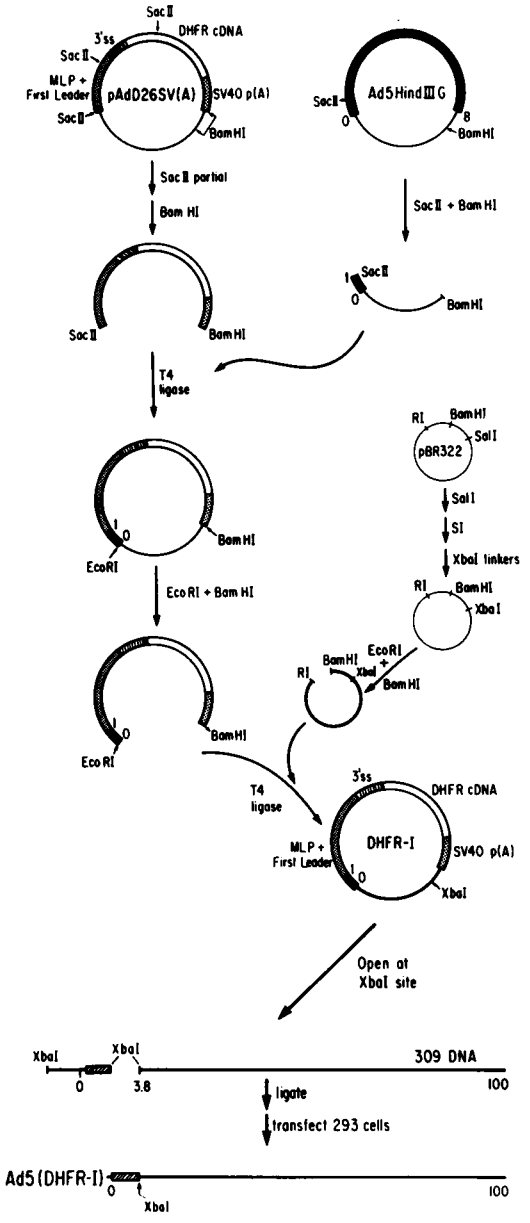

Figure 1: Generation of the recombinant Ad5 (DHFR-I). The structure for pAdD26SV (A) is presented elsewhere (5). Ad5 HindIIIG (0-8 mu) contains the entire left-end sequence of Ad5 DNA, adjoined to an Eco RI linker at 0 mu (8). The sequences upstream from the Ad5 0 mu terminus are removed during the generation of viral progeny (3).

positioned upstream from mouse DHFR cDNA. A polyadenylation signal from the early transcriptional unit of SV40 [the BclI (0.19 mu) to BamHI (0.14 mu) fragment] was inserted downstream from the coding sequences. The details of expression of this modular gene in CHO cells have been described (5). The first AUG in the spliced mRNA transcribed from this construct is the initiation codon for DHFR. Conversion of the pBR322 DNA SalI site to an XbaI site (Fig. 1) allowed the generation of Ad5 (DHFR-I) virus stocks, after ligation of the Ad5 (DHFR-I) plasmid to a purified fragment of XbaIA 309 DNA (3.8-100 mu) and transfection onto 293 cells.

Viral DNA isolated from Ad5 (DHFR-I) infected 293 cells was characterized by restriction endonuclease digestion, and the changes in fragmentation patterns expected from substituting the DHFR gene into the EIa region were observed (data not shown). In addition, the Ad5 (DHFR-I) recombinant exhibited a greater than 10^4-fold specificity for growth on 293 cells when compared with HeLa cells, while wt Ad5 grew equally well on both cell lines. Although the Ad2 MLP and adjacent sequences are homologous to the endogenous Ad5 MLP at mu 16.5, no evidence for deleted viral DNAs generated by recombination was observed. The Ad5 (DHFR-I) recombinant is 102.6 mu in length.

Analysis of polyA$^+$ RNAs produced in Ad5 (DHFR-I) infected 293 cells

A transcript originating from the Ad2 MLP, and using the appropriate processing signals built into the modular gene, should be 1.2 kb in length, and have the structure illustrated in Fig. 2. A DNA probe end-labeled at an AccI site within the DHFR cDNA, and extending to the left end of the Ad5 genome (at 0 mu), was used in S1 analysis of RNAs produced early or late after infection with wt Ad5 (309) or Ad5 (DHFR-I) virus (Fig. 2A). DHFR RNA was detected only in recombinant-infected 293 cells, and only late in infection. Even with longer exposures, no trace of DHFR RNA was observed with RNAs from early-infected cells. Two RNA species were observed, consistent with a spliced and unspliced form, both having initiated at the Ad2 MLP. The sizes of these mRNAs, as well as the shift in migration of bands in denaturing and native gels, is consistent with these two mRNAs differing by the excission of the intron.

The RNAs from Ad5 (309) or Ad5 (DHFR-I) infected 293 cells were also analyzed by Northern blotting and hybridization with probes containing either DHFR sequences, EIb sequences, or sequences from two late families of RNA (L4 and L5) (Fig. 2B). Again, RNA complementary to DHFR was observed only with Ad5 (DHFR-I)-infected cells, and only late in infection. The most prominent

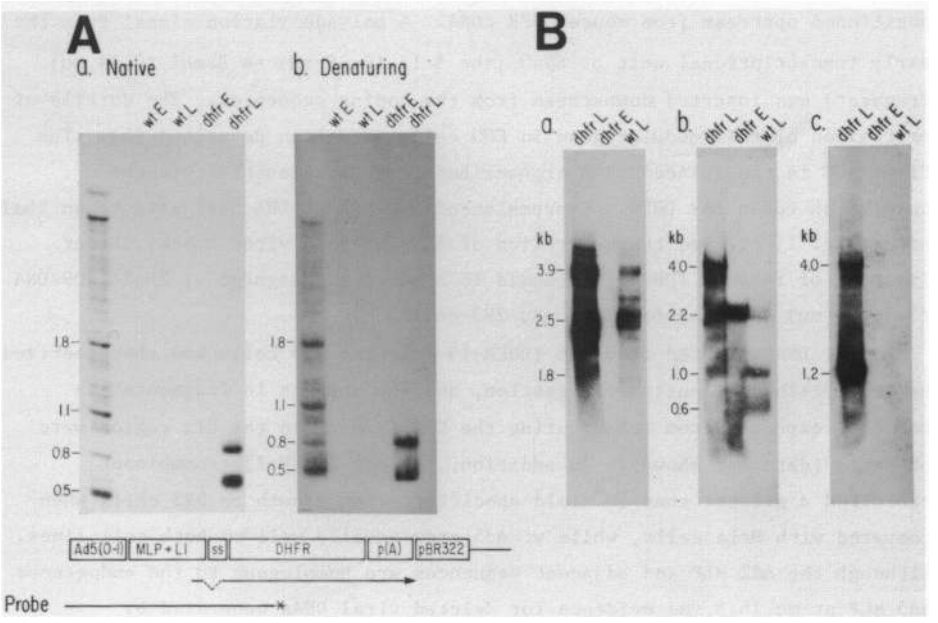

Figure 2: Analysis of RNA from viral-infected cells.

A. Early (7 hr) or late (22 hr) polyA+ RNAs, analyzed by S1 nuclease, were electrophoresed on a 1.4% neutral agarose gel (a) or glyoxylated and resolved on a denaturing gel (19) (b). The probe was 5' end-labeled at an AccI site within the DHFR cDNA sequences. The Ad2 first leader, indicated in the above diagram as L1, includes sequences up to the Ad5 HindIII site at mu 16.5, and thus includes the 5' splice site for this gene.

B. Northern analysis of polyA+ RNA (50 ng for late RNAs, 500 ng for early RNAs) was hybridized with the following probes: a) an Ad5 HindIIIB (73-89 mu), which detects 100K (4 kb), 33K (2.5 kb), pVIII and fiber (1.8 kb) mRNAs; b) EIb sequences between Ad5 3.8 and 11.1 mu; c) DHFR cDNA sequences, an SV40-pBR322 vector (20).

band complementary to DHFR was 1.2 kb in length and, coupled with the S1 data (Fig. 2A), was consistent with polyadenylation of the mRNA at the SV40 early pA+ site. The level of DHFR RNA was compared with Ad RNAs (100K, 33K and pVIII from L4 and fiber (IV) from L5) present in abundance late in infection. In general, the levels of 1.2 kb DHFR RNA were comparable, if not even higher, than mRNAs from L4 and L5, which encode abundant virion proteins (Fig. 2B).

In addition to the anticipated 1.2 kb DHFR RNA, a second, larger (4 kb) RNA containing DHFR sequences was also detected (Fig. 2B). This RNA was present at approximately 1/20th the level of the 1.2 kb species. A 4 kb RNA unique to Ad5 (DHFR-I) late-infected 293 cells was also detected by

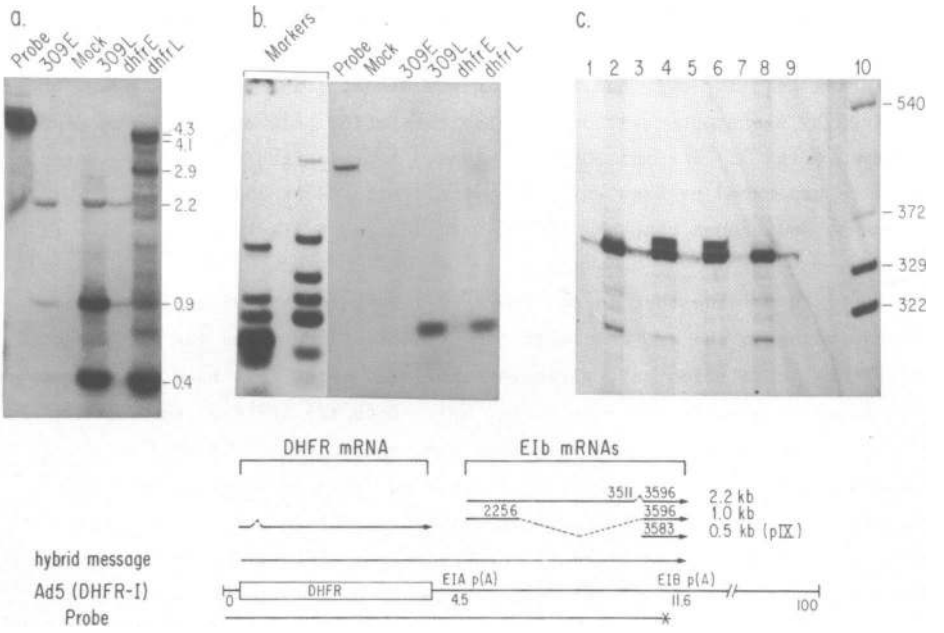

Figure 3: S1-nuclease analysis of EIb and DHFR/EIb RNA. The probe is a 4.8 kb DNA fragment end-labeled at the 11.1 mu SmaI site. For clarity the nucleotide numbers for the EIb splice sites are given for wt DNA, although the Ad5 (DHFR-I) genome is 2.6 mu longer. DNA samples were electrophoresed (a) in a 1.2% neutral agarose gel, b) through a denaturing gel, or c) in a sequencing gel (21). The RNAs in c were from early or late Ad5(DHFR-I)-infected cells (1, 2); early or late Ad5(309) infected cells (3, 4); early or late Ad5(312) (10) infected cells (5, 6); early or late Ad5(309) and Ad5(DHFR-I)-coinfected cells (7, 8); mock-infected cells (9).

hybridization with EIb sequences. This RNA could result from the failure to polyadenylate at the early SV40 site, with polyadenylation instead at the downstream EIb site.

In order to further analyze the structure of RNAs from the EIb region, S1 analysis was performed using a probe end-labeled within the second exon of the EIb region (Fig. 3). This probe should detect all three EIb mRNAs normally made during wt Ad infection, as well as potential DHFR/EIb hybrid RNAs. As can be seen in Fig. 3a, three large (2.9, 4.1 and 4.3 kb) RNAs complementary to the probe were only observed late after infection of 293 cells with Ad5 (DHFR-I). The longest of these RNAs was the same length as the RNA detected by the DHFR probe in the Northern analysis, suggesting that this mRNA was polyadenylated at the downstream EIb site. The variety of high

MW RNAs is most likely due to utilization of the EIb early and late splice signals. When the products of the same S1 experiment were resolved on a denaturing gel, the only band observed was 0.4 kb (Fig. 3b). Moreover, when resolved by electrophoresis on a higher resolution gel, only one band was observed (Fig. 3c) for Ad5 (DHFR-I) mRNAs. A band of this size (347 bases) would be generated by RNAs spliced at the EIb 3' SS at 3596 bp (see Fig. 3, diagram), and argues that the hybrid DHFR RNAs most likely use the EIb splice signals.

It is surprising that in Ad5 (DHFR-I) late infected 293 cells, a band corresponding to the expected size for pIX mRNA (360 bases) was not detected in the high resolution gel, although normal levels of this band were observed with RNA from wt infected cells (Fig. 3c). Both Ad5 (DHFR-I) and Ad5 (309)-late infected RNAs contained the same amount of 0.4 kb mRNA on a neutral gel (Fig. 3a), the size that would be expected only for pIX mRNA. Moreover, Ad5(DHFR-I)-infected cells clearly contained pIX-encoding RNA, since pIX protein was observed in infected cells (following section). The pIX message in Ad5(DHFR-I)-infected cells must have lost the sequences between the normal pIX initiation site at 3583 and the EIb 3' SS at 3596 base pairs (22), resulting in the comigration of pIX mRNA with the other EIb mRNAs on the higher resolution gel. The transcription of pIX mRNA must therefore have either initiated at 3596 base pairs or, more likely, initiated at upstream sequences, which spliced onto the EIb 3' SS.

The small difference in size observed for pIX mRNA analyzed on a native or denaturing gel (Fig. 3) raised the possibility that the putative upstream sequences spliced onto the EIb 3' SS at 3596 base pairs came from the Ad2 first leader in the DHFR modular gene. In order to test this, wt and DHFR-I-late infected RNAs were hybridized to an end-labeled probe from the second EIb exon, and were reverse transcribed (Fig. 4). The extended product for wt RNA (95 bases) was the size expected for the pIX mRNA structure. This size product was not observed for RNA from DHFR-I late infected cells. Instead, a higher MW extended product (122 bases) was observed, which was the correct size for that predicted for the Ad2 first leader splicing onto the EIb 3' SS. In order to demonstrate the presence of leader sequences on this extended product, the DNA was excised from the gel, hybridized to an M13 clone containing complementary sequences to the Ad2 leader, and digested with Pvu II. This endonuclease cleaves 33 nucleotides from the first leader cap site. When analyzed on a sequencing gel (Fig. 4), the predicted shift in fragment size for Pvu II-cleaved hybrid DNA was observed, only after

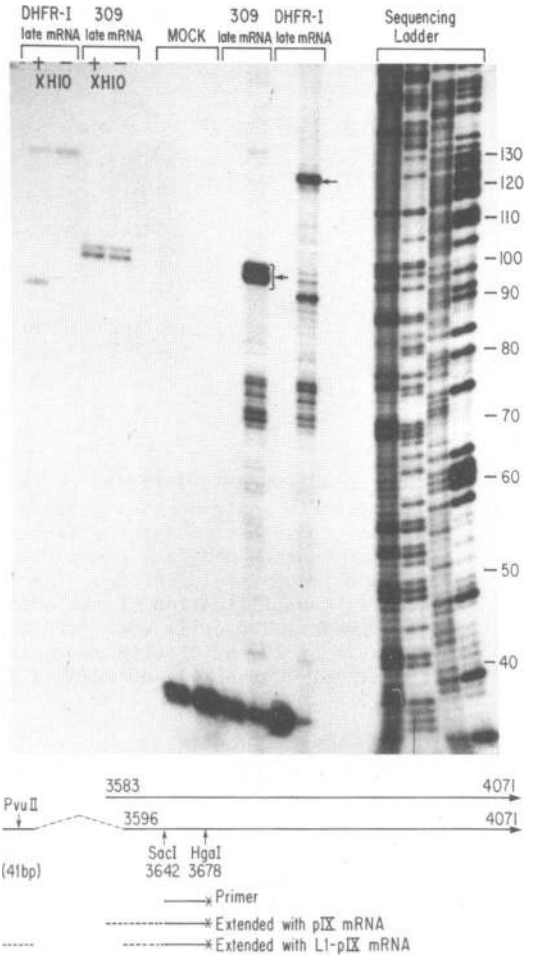

Figure 4: Primer extension of pIX mRNA from Ad5 (DHFR-I) or Ad5 (309)-infected 293 cells. A single-stranded 36 base DNA primer (2 ng), end-labeled at the HgaI site, was hybridized with polyA+ mRNA from Ad5(309) late or Ad(DHFR-I) late or mock-infected 293 cells, reverse transcribed, and electrophoresed, as described elsewhere (23). The extended products indicated by the arrows were isolated from the gel and hybridized with or without pXH10 DNA, an M13 plasmid containing the Ad5 sequences between 15.5-170 mu. Each sample was then digested with PvuII and subsequently electrophoresed through an 8% sequencing gel.

hybridization with the M13 clone.

The suppression of normal pIX mRNA transcription appeared to be trans-acting. In 293 cells coinfected with equal amounts of Ad5 (309) and Ad5 (DHFR-I), wt pIX mRNA was not observed (Fig. 3c).

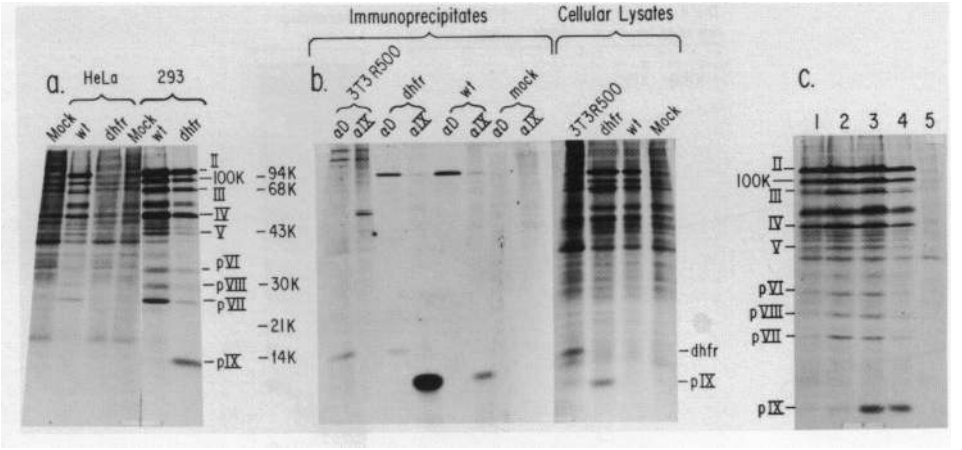

Figure 5: Protein expression in Ad5(DHFR-I)-infected cells. a) Biorad low MW markers were included, and protein assignments are based on migration relative to these. b) The film exposure time for the immunoprecipitates was three times that of the cellular lysates and five times the amount of cellular lysate was used in each immunoprecipitation. The mouse cell line 3T3R500 has a several-hundred fold amplification of the endogenous DHFR gene (24). c) In coinfection experiments, 293 cells were infected with 1) Ad5(309) or 2) Ad5(d1312) at mois of 20, or 3) with an equal mixture of Ad5(DHFR-I) and Ad5(309) (total moi=40), or 4) Ad5(DHFR-I) at a moi of 20, or 5) mock-infected.

The 1 kb EIb mRNA was found to be dramatically decreased late in Ad5 (DHFR-I)-infected 293 cells (Fig. 3a). When compared with wt late-infected cells, the level of the EIb 1.0 kb mRNA was approximately 10-fold lower. The level of 2.2 kb mRNA in Ad5(DHFR-I)-late infected cells could not be compared with wt, since it is normally present in low abundance, and was difficult to detect above background in the S1 analysis. Wild type levels of both the EIb 1 kb and the 2.2 kb mRNA were present during the early phase of infection (Fig. 3a).

Protein products produced in Ad5 (DHFR-I)-infected 293 cells

To determine if protein synthesis was specified by the DHFR mRNA, cells infected with Ad5 (309) or Ad5 (DHFR-I) were labeled with ^{35}S methionine early or late after infection, and the proteins were resolved by electrophoresis in an SDS-acrylamide gel (Fig. 5). Both viruses gave patterns indistinguishable from mock-infected cells early (6 hr) in infection (data not shown). Ad5 (DHFR-I) is defective for replication in HeLa cells, and did not stimulate the synthesis of late proteins in this cell line (Fig. 5a). Both Ad5 (309) and Ad5 (DHFR-I) infection of 293 cells resulted in production of late viral proteins, with one striking difference: the latter

vastly overproduced a 15 k dalton protein (Fig. 5). This protein was identified as the virion component pIX by immunoprecipitation with monoclonal α-pIX antisera, and was present at 40-fold higher levels than that observed after Ad5 (309) infection. Wild type Ad5 does not inhibit pIX overproduction: in mixing experiments where Ad5 (309) and Ad5 (DHFR-I) were coinfected onto 293 cells, the same level of pIX protein expression was observed as with Ad5 (DHFR-I) alone (Fig. 5c).

The production of DHFR, a 20K protein, was not observed in total cell lysates of Ad5 (DHFR-I) late-infected cells (Fig. 5a). A small amount of DHFR was observed after immunoprecipitation of the Ad5 (DHFR-I) late-infected 293 cellular extract with α-DHFR (Fig. 5b), but this quantity certainly did not reflect the large amount of DHFR RNA present late in infection (Fig. 2). The low level of protein observed was most likely not due to protein instability, since DHFR was also not observed in lysates from 293 cells pulse-labeled for 5' late in infection (data not shown).

The DHFR protein encoded by the Ad5 (DHFR-I) recombinant was determined to be functional by the criterion of conferring methotrexate (MTX) resistance to DHFR deficient Chinese hamster ovary (CHO) cells [DUKX-B1 cells (7)]. These cells were infected with Ad5 (DHFR-I) at several multiplicities, and two days post-infection were split into selective α- media containing 0, 0.005 or 0.02 uM MTX. After two weeks, MTX resistant colonies were stained and counted. The level of transformation of these CHO cells was high (1-2%). For example, infection of $5x10^5$ cells with 100 PFU of Ad5 (DHFR-1) gave 2300, 35 and 5 colonies on plates containing 0, 0.005 and 0.2 uM of MTX, respectively. No colonies were observed after infection with Ad5 (309). These results clearly demonstrate the functional nature of the DHFR encoded by Ad5 (DHFR-I).

In vitro translation of Ad5 (309) and Ad5 (DHFR-I) generated RNAs

As shown above, the high level of synthesis of DHFR mRNA did not yield comparably high levels of DHFR protein synthesis. To test whether Ad5 (DHFR-I) late infected mRNA was competent for translation, it was used to program rabbit reticulocyte lysates, (Fig. 6). High levels of DHFR protein synthesis were observed *in vitro*, only for RNAs from Ad5 (DHFR-I)-infected 293 cells at the late stage of infection. When these protein levels were quantitated, the amount of DHFR synthesis was found to be comparable to that of other abundant late Ad proteins, such as fiber. When the difference in met contained in fiber and DHFR (26, R. J. Roberts, personal comm.) was used to correct for the relative intensities by ^{35}S-met labeling, the ratio of DHFR to fiber synthesis was 1.25.

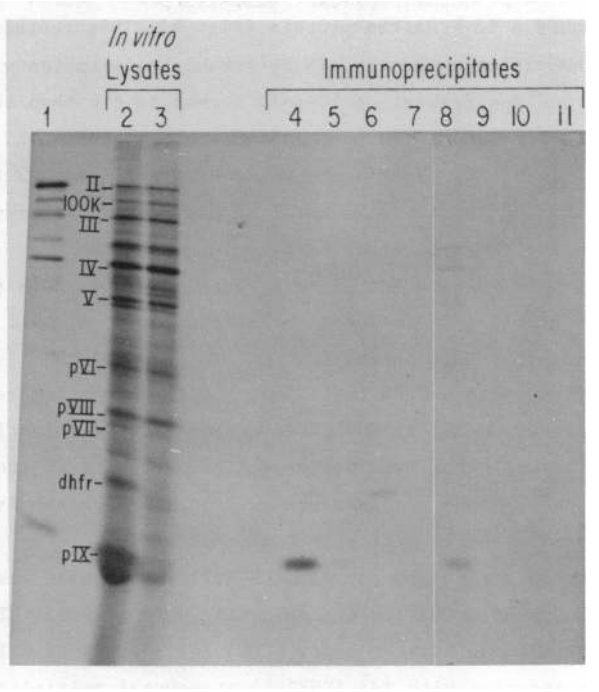

Figure 6: In vitro translation of Ad5(DHFR-I) and Ad5(309) mRNAs. PolyA$^+$ RNAs (50 ng) from Ad5(DHFR-I) (lane 2) or Ad5(309) (lane 3) late infected 293 cells were subjected to in vitro translation in a met rabbit reticulocyte lysate, purchased from BRL (25). The final reaction mix (25 ul) was 0.1 M in KAC, 1 mM in $Mg(Ac)_2$, and contained 50 ng of polyA+ RNA and 50 uCi (1000 Ci/mmol) of translation grade ^{35}S-met. After an incubation at 30^{o}C for 1 hr, 5 ul of each aliquot was electrophoresed as described in Materials and Methods. ^{35}S-labeled in vivo Ad5(DHFR-I) late proteins are shown in lane 1. Each immunoprecipitation used 1/3 the material shown for the lysate, and each immunoprecipitate supernatant was subjected to a second cycle of precipitation with antisera, in order to test for quantitative recovery. The in vitro-translation products from Ad5(DHFR-I) late RNA were precipitated with αpIX (4,5) or αDHFR (6,7), and from Ad5 (309) late RNA, precipitated with αpIX (8,9) or αDHFR (10,11).

A second difference observed with RNAs from Ad5 (309) and Ad5 (DHFR-I)-late infected 293 cells was in the amount of pIX synthesized in vitro. Translation of Ad5 (DHFR-I) RNA resulted in a dark band around the position of pIX, which migrates just behind globin (Fig. 6). After immunoprecipitation with α-pIX monoclonal antibody, however, it could clearly be seen that the level of pIX was higher (10-fold) when programmed with Ad5 (DHFR-I) RNA than with Ad5 (309) RNA (Fig. 6). The relative levels of other late RNAs (e.g. hexon, 100K, penton), though, were the same for Ad5 (309) and Ad5 (DHFR-I).

Ad5(DHFR-I) late-infected RNA contained multiple messages encoding both pIX and DHFR (Figs. 2-4). To distinguish which RNAs were being used to program translation in vitro, Ad5 (DHFR-I)-late infected polyA$^+$ RNA was fractionated on an agarose gel containing methyl mercury, and each fraction was subjected to in vitro translation and immunoprecipitation with antisera against pIX or DHFR (data not shown). The RNA programming DHFR synthesis was determined to be 1.2 kb in length, while the size of the RNA generating pIX protein was 0.5 kb. No DHFR or pIX protein synthesis was programmed by the high molecular weight DHFR-EIb mRNAs.

DISCUSSION

We have begun studying the expression of genes in Ad5 vectors by substituting the EIa region with a modular DHFR gene. The inserted transcriptional unit is composed of these segments: the major late promoter (MLP) and first leader with 5' SS, and a 3' SS positioned before the mouse DHFR cDNA, followed by the polyadenylation site from the early region of SV40 (see Fig. 1). A 1.2 kb mRNA was synthesized from the DHFR gene in Ad5(DHFR-I)-infected 293 cells, indicating that the splicing and polyadenylation signals were utilized. At late times, the transcriptional efficiency of the inserted Ad2 MLP was roughly the same as that of the endogenous MLP at mu 16.5. Thus approximately equal amounts of DHFR mRNA and late mRNAs, including the abundant 100K mRNA, were observed. Transcription of DHFR mRNA from the inserted MLP was only observed late in Ad infection even under conditions where other early messages (i.e. EIb) were detected. This is somewhat surprising, since the MLP is known to function early in infection, at levels slightly less than those of other early promoters (27). DHFR mRNAs were also not detected in Ad5 (DHFR-I) infected HeLa cells, where the absence of an EIa product precluded replication and late gene expression.

Although high levels of DHFR mRNAs were present in recombinant-infected 293 cells at late times, very little DHFR protein synthesis was detected by pulse labeling with [^{35}S]-met and immunoprecipitation with α-DHFR anti sera. The DHFR gene in Ad5 (DHFR-I) is biologically active, since this virus efficiently (1-2%) transformed DHFR(-) CHO cells to DHFR (+) phenotype. To test whether the DHFR mRNA structure in late infected cells precluded translation, this RNA was used to program translation in a cell-free reticulocyte lysate. High amounts of DHFR protein were observed, reflecting the high concentration of DHFR mRNA present in late-infected RNA (Fig. 6). The relative in vitro translation efficiency of DHFR mRNAs and other late Ad5

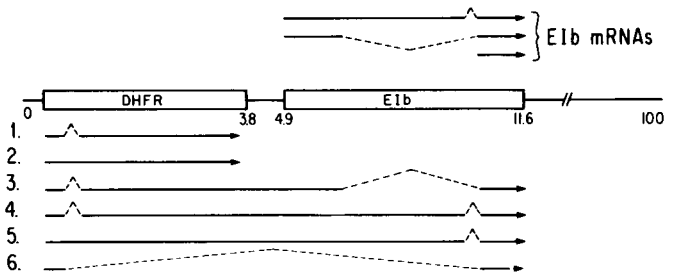

Figure 7: Novel RNAs observed after infection of 293 cells with Ad5(DHFR-I). Ad5(DHFR-I) and wt Ad5(309) DNA have identical sequences from 3.8 to 100 mu. For clarity, the EIb unit is identified using wt mu assignments, although Ad5(DHFR-I) DNA is 102.6 mu in length.

mRNAs were comparable. This is in stark contrast to the in vivo results, and probably reflects the pronounced translation regulation imposed by adenovirus in late infected cells. The DHFR mRNAs contained only the first leader of the tripartite set found on late adenoviral mRNAs. The additional sequences from the tripartite leader are probably important for efficient translation at the late stage in vivo, and may be required for mediation of the recently recognized translation control by the small virus associated (V.A.) RNAs (28). Previous studies have also described Ad recombinants with 2/3 of the first leader positioned upstream from EIa or early SV40 (29). These recombinants also produced high levels of RNA but exhibited poor translatability in vivo.

Expression of the EIb unit, downstream from the DHFR gene in Ad5 (DHFR-I), was altered in late infected 293 cells. First, transcription from the EIb and the pIX promoters was suppressed. This may be due to the presence of the strong upstream MLP. Second, two types of novel, EIb-containing mRNAs were produced. One set, 3-4 kb in length (Fig. 7, numbers 3-5), contained both DHFR and EIb sequences, and apparently varied from each other by alternative splicing patterns, using the EIb splice sites active in vivo. The 3-4 kb RNAs were polyadenylated at the EIb site, even though the transcript contained three other pA signals: in the DHFR cDNA, in the early polyA site of SV40 and in the EIa site at 4.5 mu. Other results have suggested that the DHFR cDNA and EIa polyadenylation signals may be inefficient (5). The other type of novel spliced mRNA, LI-IX, was 0.5 kb in length, and resulted from splicing of the first leader to the 3' SS at position 3596 in EIb (Fig. 7, number 6). The level of expression of this novel RNA was comparable to that of the pIX mRNA in a w.t. Ad5 late

TABLE I

mRNA	Sequence
pIX	CAGCCGCCGCCGCCAUG
hexon	GAGCCGCCGU-$(N)_{15}$-AUG
pIX	CAGCCGCCGCCGCCAUG
23K	CUGCCGCCAUG
pIX	CAGCCGCCGCCGCCAUG
pVIII	CAG CGCCAUUAUG
pIX	CAGCCGCCGCCGCCAUG
L2	UCGGCCUCCGA
pIX	CAGCCGCCGCCGCCAUG
L3	UCCGCCACCG
pIX	CAGCCGCCGCCGCCAUG
LY	ACCACCACCACCACC

Sequence homology of pIX mRNA with late leader messages. The references for the sequence determination of each RNA are as follows: pIX (22), hexon (30), 23K (31), pVIII (32), L2 and L3 (30, 33), and LY (33). The initiation AUG is underlined.

infection.

The efficiency of in vivo translation of the pIX-containing mRNA from Ad (DHFR-I) infected cells was 40-fold higher than that of the normal wt pIX mRNA (Fig. 5). The mRNA which is so efficiently translated appears to be the LI-IX mRNA. Although the novel 3-4 kb hybrid mRNAs encode pIX, translation of this polypeptide would require initiation at the 6th AUG in these messages. Furthermore, with similar constructions to the Ad5(DHFR-I) described here, where the 3-4 kb hybrid mRNAs are detected in infected cells, pIX overproduction is not observed (our unpublished data). The efficiency of in vitro translation of this LI-IX mRNA was 10-fold higher than that of the normal pIX mRNA. This is surprising and suggests that in the absence of viral regulation, the sequence structure of the L1-IX mRNA is more suited to initiation of translation than that of the pIX mRNA. However, this intrinsic increase in efficient translation probably does not totally explain the overproduction of polypeptide IX from the L1-IX mRNA in late infected cells, since translation efficiency is strictly regulated at this time (28). Thus, the configuration of sequences in the L1-IX mRNA must also be efficiently recognized by the virally induced translation state. Part of this recognition could be contributed by the L1 leader sequence, but not all, as indicated by the inefficient synthesis of DHFR protein in these cells. A

w.t. pIX mRNA must contain some sequence elements that permit translation during late infection. The high translation efficiency of the L1-IX mRNA might be explained by the combining of elements from the L1 leader with that of the normal pIX mRNA.

It might be anticipated that some of the elements in the L1-IX mRNA contributed by the pIX sequences are also in the second and third exons in the tripartite leader of late mRNAs. In the novel L1-IX transcript, the 14 bases derived from the pIX 5'-noncoding sequences contain a striking 4-fold repeat of the sequence, GCC (Table I). In a comparison with individual parts of the tripartite leader, a partially homologous region was observed with both the second (7 base homology with one mismatch) and third (9 base homology with one mismatch) leaders (30, 33). Interestingly, there is a perfect homology of eight of the pIX-specified nucleotides with part of the hexon 5'-noncoding sequences (Table I; 30). In addition, extensive homology is also observed with two other late messages: pVIII mRNA and the L3 mRNA which encodes a 23K protein (31, 32). The "Y" leader, which contains a 15 nucleotide stretch consisting of the repeat ACC, is also homologous (33). The corresponding pIX 5' noncoding sequence in Ad7 (34) is substantially homologous to that in Ad2, but there is no obvious sequence similarity with the 5'-noncoding pIX sequences of Ad12 (35). It should be noted that a tabulation of sequences upstream of AUGs used for initiation of translation of cellular proteins produces a consensus sequence of $CC\begin{matrix}A\\G\end{matrix}CCAUG$ (36). The GCC repeat preceeding the initiation AUG on IX mRNA is identical to the consensus sequence, suggesting that this mRNA may inherently be an unusually good substrate for translation.

ACKNOWLEDGMENTS

We wish to thank P. Jat and M. Kozak for constructive comments relative to this manuscript. M. Samuels, R. Kaufman and M. Horowitz provided plasmid DNAs, and C. Cepko provided αpIX monoclonal α serum, all of which were gratefully appreciated. We also want to thank M. Siafaca for her careful preparation of this manuscript.

K.L.B. is the recipient of a Charles A. King Trust fellowship from The Medical Foundation, Inc. This work was supported by grants from National Science Foundation [PCM-7823230 (currently PCM-8200309)], from National Institutes of Health (No. PO1-CA26717) to P.A.S. and partially from Center for Cancer Biology at MIT (Core) grant (No. NIH-PO1-CA14051).

REFERENCES

1. Graham, F. L., Smiley, J., Russell, W. C. and Nairn, R. (1977) J. Gen. Virol. 36, 59-72.
2. Grodzicker, T. and Klessig, D. F. (1980) Cell 21, 453-463.
3. Berkner, K. L. and Sharp, P. A. (1983) Nucl. Acids Res. 11, 6003-6020.
4. Flint, S. J. (1982) Biochim. Biophys. Acta 651, 175-208.
5. Kaufman, R. J. and Sharp, P. A. (1982) Mol. Cell. Biol. 2, 1304-1319.
6. Kaufman, R. J. and Sharp, P. A. (1982) J. Mol. Biol. 159, 601-621.
7. Chasin, L. A. and Urlaub, G. (1980) Proc. Natl. Acad. Sci. USA 77, 4216-4220.
8. Bolivar, F., Rodriguez, R. C., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. and Falkow, S. (1977) Gene 2, 95-113.
9. Vogelstein, B. and Gillespie, D. (1979) Proc. Natl. Acad. Sci. USA 76, 615-619.
10. Jones, N. and Shenk, T. (1979) Cell 17, 683-689.
11. Graham, F. L. and van der Eb, A. J. (1973) Virol. 52, 456-467.
12. Seed, B. and Goldberg, D., manuscript in preparation.
13. Berk, A. J. and Sharp, P. A. (1977) Cell 12, 721-732.
14. Laemmli, U. K. (1970) Nature (London) 227, 680-685.
15. Kessler, S. W. (1975) J. Immunol. 115, 1617-1624.
16. Cepko, C. L., Changelian, P. S. and Sharp, P. A. (1981) Virology 110, 385-401.
17. Hammarskjöld, M. L. and Winberg, G. (1980) Cell 20, 787-795.
18. Tibbetts, C. (1977) Cell 12, 243-249.
19. McMaster, G. K. and Cormichael, G. G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
20. Horowitz, M., Cepko, C. L. and Sharp, P. A. (1983) J. Mol. Applied Genetics 2, 147-159.
21. Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
22. Aleström, P., Akusjärvi, G., Perricaudet, M., Mathews, M. B., Klessig, D. F. and Pettersson, U. Cell 19, 671-681.
23. Jat, P. S. (1982), thesis, University of London.
24. Kellems, R. E., Morhenn, V. B., Pfendt, E. A., Alt, F. W. and Schimke, R. T. (1979) J. Biol. Chem. 254, 309-318.
25. Pelham, H. R. and Jackson, R. J. (1976) Eur. J. Biochem. 67, 247-256.
26. Stone, D. and Phillips, A. W. (1977) FEBS Letters 74, 85-87.
27. Lewis, J. B. and Mathews, M. B. (1980) Cell 21, 303-313.
28. Thimmappaya, B., Weinberger, C., Schneider, R. J. and Shenk, T. (1982) Cell 31, 543-551.
29. Solnick, D. (1981) Cell 24, 135-144.
30. Akusjärvi, G. and Pettersson, U. (1979) Cell 16, 841-850.
31. Akusjärvi, G., Zabielski, J., Perricaudet, M. and Pettersson, U. (1981) Nucl. Acids Res. 9, 1-7.
32. Galibert, F., Herisse, J. and Courtois, G. (1979) Gene 6, 1-22.
33. Zain, S., Sambrook, J., Roberts, R. J., Keller, W., Fried, M. and Dunn, A. R. (1979) Cell 16, 851-861.
34. Dijkema, R., Maat, J., Dekker, B. M. M., van Ormondt, H. and Boyer, H. W. (1981) Gene 13, 375-385.
35. Tooze, J., ed. (1981) DNA Tumor Viruses, p. 1028.
36. Kozak, M. (1984) Nucl. Acids Res., in press.