Nucleic Acids Research, Vol. 19, No. 16 4485–4490

Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus

Randal J.Kaufman, Monique V.Davies, Louise C.Wasley, and Donna Michnick
Genetics Institute, 87 Cambridge Park Drive, Cambridge, MA 02140, USA

Received May 1, 1991; Revised and Accepted July 22, 1991

ABSTRACT

Dicistronic mRNA expression vectors efficiently translate a 5′ open reading frame (ORF) and contain a selectable marker within the 3′ end which is inefficiently translated. In these vectors, the efficiency of translation of the selectable 3′ ORF is reduced approximately 100-fold and is highly dependent on the particular sequences inserted into the 5′ cloning site. Upon selection for expression of the selection marker gene product, deletions within the 5′ ORF occur to yield more efficient translation of the selectable marker. We have generated improved dicistronic mRNA expression vectors by utilization of a putative internal ribosomal entry site isolated from encephalomyocarditis (EMC) virus. Insertion of the EMC virus leader sequence upstream of an ORF encoding either a wildtype or methotrexate resistant dihydrofolate reductase (DHFR) reduces DHFR translation up to 10-fold in a monocistronic DHFR expression vector. However, insertion of another ORF upstream of the EMC leader to produce a dicistronic mRNA does not further reduce DHFR translation. In the presence of the EMC virus leader, DHFR translation is not dependent on sequences inserted into the 5′ end of the mRNA. We demonstrate that stable high level expression of inserted cDNAs may be rapidly achieved by selection for methotrexate resistance in DHFR deficient as well as DHFR containing cells. In contrast to previously described dicistronic expression vectors, these new vectors do not undergo rearrangement or deletion upon selection for amplification by propagation in increasing concentrations of methotrexate. The explanation may be either that the EMC virus leader sequence allows internal initiation of translation or that cryptic splice sites in the EMC virus sequence mediate production of monocistronic mRNAs. These vectors may be generally useful to rapidly obtain high level expression of cDNA genes in mammalian cells.

INTRODUCTION

The scanning model for translation initiation postulates that a 40S ribosomal subunit binds to the 5′ capped end of the mRNA and migrates in the 3′ direction until it encounters the first AUG triplet present in an appropriate context to serve as an initiator codon. Subsequently the 60S ribosomal subunit joins and polypeptide chain synthesis begins (1). Although a substantial amount of evidence supports a 'scanning' model for translation initiation, there is now some evidence that a few naturally occuring mRNAs have the capability to initiate at internal AUG codons (for review see 2). Although no functional polycistronic cellular mRNAs with non-overlapping open reading frames in higher eukaryotes have been identified to date, studies using experimental constructs have shown that initiation at internal AUG codons occurs (3,4). Thus it is possible to translate two or more proteins from non-overlapping open reading frames (ORFs) within a single mRNA. In experimental constructs containing dicistronic mRNAs, initiation at the downstream ORF is generally 100-fold less efficient than initiation at the upstream AUG of the first ORF (4,5). To account for the inefficient initiation at internal AUG codons, the scanning model has been modified to allow for reinitiation at AUG codons downstream from a translated ORF. The ability to select for the low level of expression due to reinitiation at an ORF within the 3′ end of the mRNA has provided an avenue to obtain high level expression using dicistronic mRNA expression vectors (4,6). In these vectors, translation of the desired ORF in the 5′ position within the mRNA is efficient while translation of the selectable marker within the 3′ position is very inefficient. Dicistronic mRNA expression vectors containing the gene of interest followed by the selectable and amplifiable marker gene dihydrofolate reductase (DHFR), can be selected in DHFR deficient CHO cells by growth in nucleoside free medium. Subsequent selection of cells by growth in low concentrations of methotrexate results in amplification of the transcription unit and increased expression of the desired gene product. However, selection for increased DHFR expression frequently results in cells that delete the first ORF in order to translate DHFR more efficiently.

One class of mRNAs for which ribosomes might be able to bind internally are those belonging to picornaviruses. The 5′ untranslated regions from poliovirus (7) and encephalomyelocarditis (EMC) virus (8) may be able to promote ribosome binding and translation initiation at internal sites within an mRNA in intact cells. In constructs which contain two ORFs, insertion of the picornavirus 5′ UTR upstream from the second, 3′ proximal, ORF yields measurable translation of that ORF independent of the presence of the first ORF. The required

sequence of EMC virus to promote internal initiation is unexpectedly large, extending from 400 to 834 nucleotides of the viral genome (9). We have derived a new set of mammalian cell expression vectors which utilize the 5′ UTR from EMCV to promote efficient internal translation intiation of selectable markers encoding murine wildtype DHFR, a MTX resistant DHFR (10), or neomycin phosphotransferase. These vectors readily yield high level expression of foreign genes in stably transfected Chinese hamster ovary cells. In addition, this approach can be used to determine whether a gene product is potentially toxic to the cell.

MATERIALS AND METHODS

Derivation of pMT21

pMT21 was derived from pMT2 (11) through the following two modifications. First, 76 bp of the 5′ untranslated region of the DHFR cDNA including a stretch of 19 G residues from G/C tailing for cDNA cloning was deleted. In this process, a XhoI site was intserted to obtain the following sequence immediately upstream from DHFR:

5′-CTGCAGGCGAGCCTGAATTCCTCGAGCCATCATG-3′
PstI Eco RI XhoI

Second, a unique ClaI site was introduced by digestion with EcoRV and XbaI, treatment with Klenow fragment of DNA polymerase I, and ligation to a ClaI linker (CATCGATG). This deletes a 250 bp segment from the adenovirus virus associated RNA (VAI) region but does not interfere with VAI RNA gene expression or function.

Derivation of pED4, pED, pED-mtx^r expression vectors

pMT2-ECAT1 (8) was digested with Eco RI and PstI resulting in a 2752 bp fragment. This fragment was digested with TaqI yielding an Eco RI-TaqI fragment of 508 bp which was purified by electrophoresis on low melting agarose gel. A 68 bp adapter was synthesized with a 5′ TaqI protruding end and a 3′ XhoI protruding end which has the following sequence:

5′-TCGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTG-
TaqI
GTTTTCCTTTGAAAAACACGATTGCTCGAG-3′
XhoI

This sequence matches the EMC virus leader sequence from nucleotide 763 to 827. It also changes the ATG at position 10 within the EMC virus leader to an ATT and is followed by a XhoI site. A three way ligation of the pMT21 Eco RI-XhoI fragment, the EMC virus EcoRI-TaqI fragment, and the 68 bp oligonucleotide adapter TaqI-XhoI adapter resulted in the vector pED4. pED4 has the authentic DHFR initiation ATG codon 15 bp downstream from the position of ATG-10 of the EMC virus. To increase the utility of pED4, pED was derived from by insertion of the following polylinker between the PstI and Eco RI cloning sites:

5′-TGCAGGTCGACTCTAGACCCGGGAATT-3′
PstI/SalI/ XbaI/SmaI/EcoRI

pED-mtx^r was derived from pED4 by site-directed oligonucleotide mutagenesis to change a leucine codon (CTA) at residue 22 to an arginine codon (CGA) as described (11).

pED4-2α and pED4-2α(48A) were obtained by inserting a 1.6 kb Eco RI fragment encoding wildtype eIF-2α or a mutant encoding a serine to alanine mutation ar residue 48, respectively, into pED4. pED-mtx^r-PACE was derived by cloning a 2.47 kb Eco RI fragment encoding PACE (12) into the Eco RI site of pED-mtx^r.

Cell Culture

COS-1 monkey cells (13) were transfected using DEAE-dextran with the addition of a chloroquin treatment as described (14). The derivation and growth of DHFR deficient CHO-GRA cells which overexpress the rat glucocorticoid receptor have previously been described (15). Rat-2 cells were propagated in Dulbecco's modified essential medium. Cells were electroporated with plasmid DNA as described (16). Briefly, subconfluent cultures of cells were trypsinized and 2×10^6 cells were resuspended in 1 ml Dulbecco's minimal essential medium supplemented with 10 mM glutamine. DNA (100 μg) was linearized by digestion with NdeI and added to the cells. This mixture was exposed to 200 V at 1250 μF (Cell ZapII, Andersen Electronics, Brookline, MA) and the cells were plated within 5 min into nonselective medium. Two days later cells were subcultured 1:15 into selective medium as described. Electroporation efficiencies are quantitated as the number of colonies per number of cells electroporated.

Analysis of protein synthesis, mRNA, and genomic DNA

Protein synthesis was monitored by labeling CHO cells or COS-1 cells with 100 μCi/ml ^{35}S-methionine (sp Activity—8000 Ci/mMole, New England Nuclear). Cell extracts were prepared by lysis in Nonidet-P 40 lysis buffer as described (11) and analyzed by SDS-polyacrylamide gel electrophoresis (17), either before or after immunoprecipitation with rabbit polyclonal anti-human eIF-2α antibody kindly provided by Dr. J.W.B.Hershey. Gels were fixed in 40% methanol and 10% acetic acid, and prepared for fluorography by treatment with EnHance (New England Nuclear), and dried. Dried gels were autoradiographed with Kodak XAR-5 film with a Du Pont Cronex Lightning-Plus screen. Protein synthesis was quantitated by visual comparison of band intensities from multiple autoradiograms of different exposure times.

Total RNA was prepared by guanidine thiocyanate extraction (18) and analyzed by Northern blot hybridization (19) following electrophoresis on formaldehyde-formamide denaturing agarose gels and transfer to nitrocellulose as described (20). Hybridization was carried out using gel-isolated restriction fragments which had been labeled with ^{32}Pα-dCTP using random priming with oligonucleotides as described by the supplier (Pharmacia Inc).

RESULTS

Derivation of EMCV Leader Containing Expression Vectors

The starting expression plasmid for derivation of the EMCV expression vectors was pMT21 (Figure 1). pMT21 is identical to pMT2 except it contains a polylinker cloning site for convenient insertion of cDNA clones (11). pMT21 contains the SV40 origin of replication and enhancer element. It utilizes the adenovirus major late promoter for transcription initiation. Contained within the 5′ end of the mRNA is the tripartite leader from adenovirus late mRNA and a small intervening sequence. The unique cloning sites are PstI, Eco RI, and XhoI. The 3′ end of the transcript contains a murine DHFR coding region and signals for cleavage and polyadenylation of mRNA from the SV40 early region. In addition, this plasmid contains the adenovirus VAI RNA gene which encodes an RNA polymerase III transcript that potentiates translation of the plasmid derived mRNA in transfected COS-1 monkey cells (11). The plasmid contains the sequences from puc18 which allow for propagation and selection for ampicillin resistance in E. coli.

Within the EMC virus leader, the 11th AUG codon from the 5′ end of the viral mRNA is the authentic initiation codon and is only 8 bases downstream from the 10th AUG codon which is in a different reading frame. Previous studies have shown that the sequence context around the authentic initiator codon may dramatically influence translation initiation in an *in vitro* reticulocyte lysate translation system (21). These results suggest that a very precise mechanism selects for initiation at the 11th AUG and that it may be essential to utilize this AUG for efficient internal translation initiation. We have evaluated the importance of the sequence context around the 11th AUG codon and this data will be presented elsewhere (Davies et al, in preparation). In the course of these experiments, we destroyed the ATG codon at position 10 and inserted an XhoI linker near the initiation codon for DHFR. The presence of the XhoI site facilitates the introduction of other coding regions behind the EMCV leader as desired. The sequence context of the final DHFR selection vector pED4 is compared to the natural EMC virus sequence in the bottom of Fig. 1.

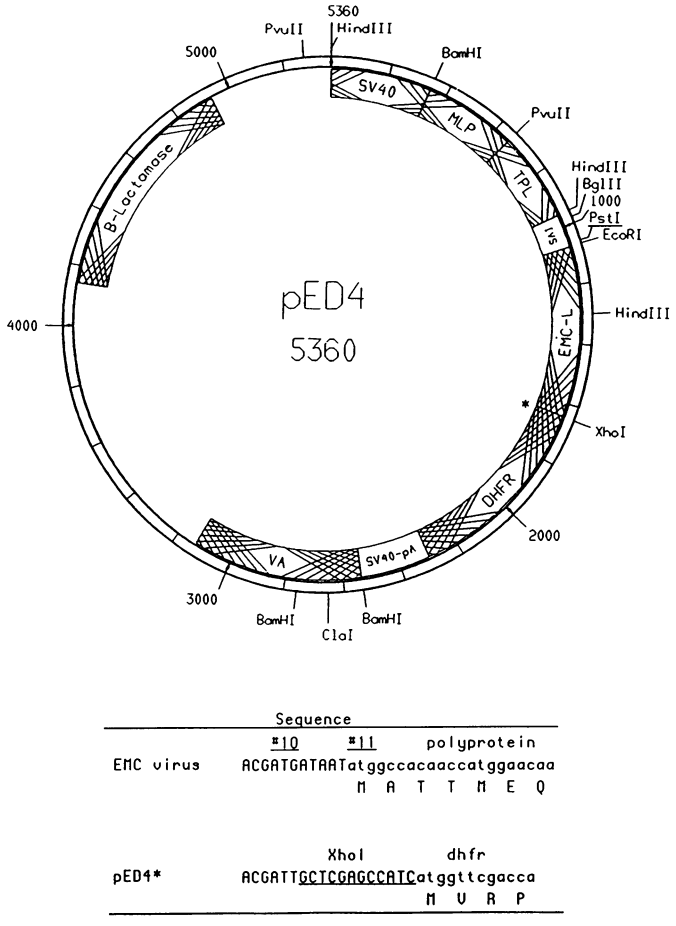

Figure 1. Map of Dicistronic mRNA Expression Vector pED4. The components of the 5360 bp pED4 expression vector in the puc18 background are indicated as follows: SV40, HindIII-PvuII fragment containing the SV40 origin of replication and enhancer element; MLP, adenovirus major late promoter from the XhoI site (15.83 m.u.) to the 5′ cap site (16.55 m.u.); TPL, 180 bp of the first two and 2/3 of the third leaders from adenovirus major late mRNAs; IVS, a hybrid intron composed of the 5′ splice site from the first leader of adenovirus major late mRNAs and a 3′ splice site from an immunoglobulin gene (26); PstI and EcoRI unique cloning sites; EMC-L, the 5′ untranslated leader from EMC virus (nucleotides 260–827); DHFR, a murine DHFR coding region; SV40-poly A, the SV40 late polyadenylation signal; VA, the adenovirus VAI RNA gene from the HpaI (28.02 m.u.) to the BalI (29.62 m.u.); and β-lactamase, a selectable gene for propagation in E. coli. Below is indicated the sequence junction of the EMC-L and DHFR (*) as compared to the context of the AUG 11 which is initiation codon for the EMC virus polyprotein. A unique XhoI restriction site is available for insertion of other coding regions to be translated from the EMC virus leader.

Expression from pED4 in COS-1 Monkey Cells

To evaluate the effect of the insertion of upstream sequences on DHFR translation from the dicistronic mRNA expressed from pED-4, a cDNA encoding human eukaryotic initiation factor 2α (eIF-2α; 11) was cloned into the unique Eco RI site immediately upstream from the EMCV leader. For a control, the eIF-2α cDNA was inserted into the starting expression vector pMT21, which does not contain the EMC virus internal ribosomal entry site.

Expression of eIF-2α and DHFR were analyzed by transient transfection of COS-1 monkey kidney cells. Synthesis of these two proteins was monitored by short ^{35}S-methionine pulse labeling of transfected cells and analysis of the total cellular protein by SDS-polyacrylamide gel electrophoresis of prepared cell extracts and fluorography (Fig. 2A). From parallel plates of transfected cells, total cellular RNA was isolated for analysis by Northern blot hybridization using a DHFR specific probe (Fig. 2B). DHFR synthesis is readily detected as a band migrating at 20 kDa in extracts from cells transfected with the monocistronic vector pMT21 (Fig. 2A; lane 2). Cells transfected with pED4 direct the synthesis of DHFR (Fig. 2A; lane 4), however DHFR synthesis is approximately 10-fold lower than that observed with pMT21. Since there is no significant difference in the DHFR mRNA level expressed from pED4 and pMT21 (Fig. 2B; lanes 1,3), the reduced DHFR synthesis results from a poorer translational efficiency.

The level of eIF-2α synthesis from pED4-2α was similar to that observed from pMT21-2α (Fig. 2A; lane 3 and 5). Thus,

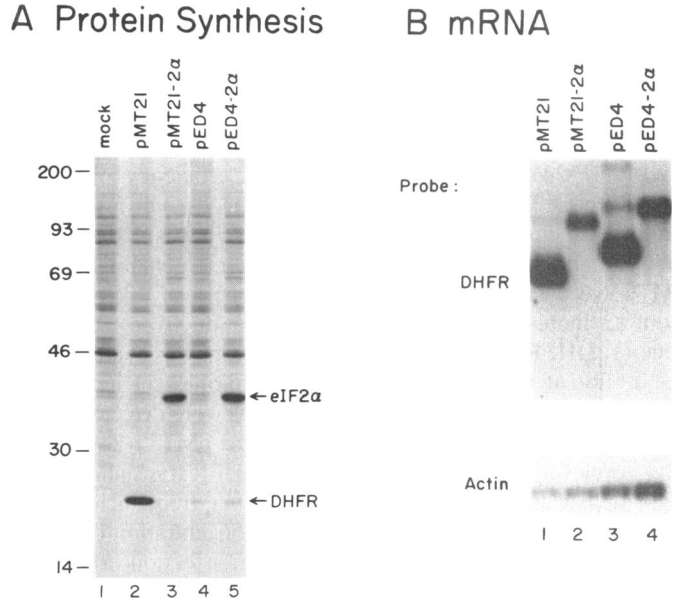

Figure 2. Expression of pED4 vectors in COS-1 cells. The indicated plasmid DNAs were transfected into COS-1 monkey kidney cells. A. At 48 hr post-transfection cells were pulse labeled for 30 min with ^{35}S-methionine and total cell extracts prepared for analysis by SDS-PAGE as described in Materials and Methods. Migration of eIF-2α and DHFR are indicated. Position of molecular weight markers (kDa) are indicated on the left. B. From parallel plates at 48 hr post-transfection, total cell RNA was harvested and analyzed by Northern blot hybridization to a DHFR and actin probes as described in Materials and Methods.

insertion of the EMC virus leader did not significantly interfere with eIF-2α translation or mRNA expression. In contrast, whereas DHFR synthesis could not be detected above background in pMT21-2α transfected cells (lane 3), synthesis of DHFR was not influenced by insertion of the eIF-2α cDNA into pED4 (Fig. 2A; lanes 4,5). Northern blot analysis of the level of mRNA produced from cells transfected in parallel was performed using a DHFR probe. DHFR hybridizes to all transcripts derived from these vectors. The monocistronic vectors pMT2 and pED4 produce similar levels of mRNA, while the dicistronic vectors pMT21-2α and pED4-2α produce several fold lower levels of mRNA. In all cases, only one species of mRNA hybridizing to DHFR was detected, indicating that DHFR translation did not result from an alternate transcript derived from the vector. mRNA levels were quantitated by hybridization to a DHFR probe and standardized to levels of actin mRNA determined by hybridization of the same blot to an actin probe. Although mRNA levels derived from pED4-2α is only 25% the level derived from pED4, DHFR synthesis from pED4-2α is similar to that observed in pED4 transfected cells. Therefore, translational efficiency of the ORF behind the EMCV leader actually increased by insertion of the eIF-2α cDNA between the adenovirus major late promoter and the EMCV leader. Similar increases in translation by insertion of RNA sequences upstream from the EMC virus internal ribosomal entry site were reported previously (8).

pED4 Dicistronic mRNA Vector Efficiently Transforms DHFR Deficient CHO Cells and Does Not Rearrange

The utility of the pED4 vector to obtain high level expression in stably transfected cells was evaluated by electroporation of pED4 and pED4-2α into DHFR deficient CHO cells. Forty-eight hrs after electroporation, cells were subcultured into selection medium lacking nucleosides in the absence and presence of increasing concentrations of methotrexate. Electroporation efficiencies were determined from the ability of plasmid DNA to transform DHFR deficient CHO cells to grow in nucleoside-free medium. Efficiencies obtained with the pED4 vectors were compared to the pMT21 vectors using an insert cDNA encoding either the wildtype eIF-2α or a serine to alanine mutant at residue 48 of eIF-2α (11). The electroporation efficiency of pED4 harboring the wildtype eIF-2α cDNA (pED4-2α) or a serine to alanine mutant of eIF-2α [(pED4-2α(48A)] was 10−40 fold greater compared to the eIF-2α cDNA insert contained in the pMT21 vector (Table I.A). This increased efficiency likely results from an increased level of DHFR translation from the pED4 vectors. DHFR expression from these vectors was evaluated by the ability of transfected cells to form colonies in increasing concentrations of methotrexate. No colonies were detected in the presence of 0.005 or 0.02 μM MTX with the pMT21-2α vector. This result reflects the low efficiency of DHFR translation from these dicistronic vectors. In contrast, higher frequencies were obtained using the pED4 vector. As the concentration of MTX was increased from 0.005 to 0.02, the transfection efficiency decreased approximately 10-fold (Table I.A). In a separate experiment qualitatively similar results were obtained comparing the pMT21 vector with the pED4 vector using a different cDNA insert.

Our previous experience with the dicistronic mRNA expression vector pMT21 demonstrated that frequently the plasmid DNA would rearrange and/or delete the inserted cDNA upon selection for DHFR. Rearrangement of the vector is associated with the generation of a new DHFR mRNA which is likely more efficiently translated to yield sufficient DHFR to survive the selection upon growth in nucleoside-free medium or upon selection for amplification by growth in methotrexate. If DHFR translation is more efficient in the pED4 vectors, then there may be less selection for rearrangements to occur. We studied whether the pED4 vector exhibits reduced rearrangement or deletion compared to the pMT21 vector by monitoring the DHFR mRNA expressed by Northern blot analysis. If mutations such as deletions or rearrangements occur to alter the mRNA transcription unit, then different cell lines should exhibit different or multiple sized mRNAs encoding DHFR. In contrast, if no rearrangements or deletions occur, then all transformants exhibit DHFR mRNAs of the same size. Northern blot hybridization to 19 independent clones transfected with pMT21-2α and selected for resistance to 0.02 μM methotrexate identifies the presence of multiple mRNA species of different sizes which hybridize to a DHFR probe (Fig. 3, bottom). Frequently, two hybridizing species are

Table I. Transfection Efficiency of pED Vectors

Cells Vector	Selection 0	0.005	0.02 (μM MTX)
A. CHO			
pMT21-2α	1×10^{-5}	$<10^{-7}$	$<10^{-7}$
pED4-2α	1.2×10^{-4}	4×10^{-5}	4×10^{-6}
pED4-2α(48A)	4×10^{-4}	5×10^{-5}	6×10^{-6}
	0.075	0.15	0.30 (μM MTX)
B. Rat-2			
No DNA	1.5×10^{-5}	$<10^{-7}$	$<10^{-7}$
pED-mtx^r-PACE	2.3×10^{-4}	2.5×10^{-5}	2×10^{-5}

The indicated plasmid DNA was electroporated into DHFR deficient CHO GRA cells (A.)(15) or into Rat-2 fibroblasts (B.). Forty-eight hrs later cells were subcultured into selective medium with increasing concentrations of MTX. Transformation efficiencies were determined as described in Materials and Methods.

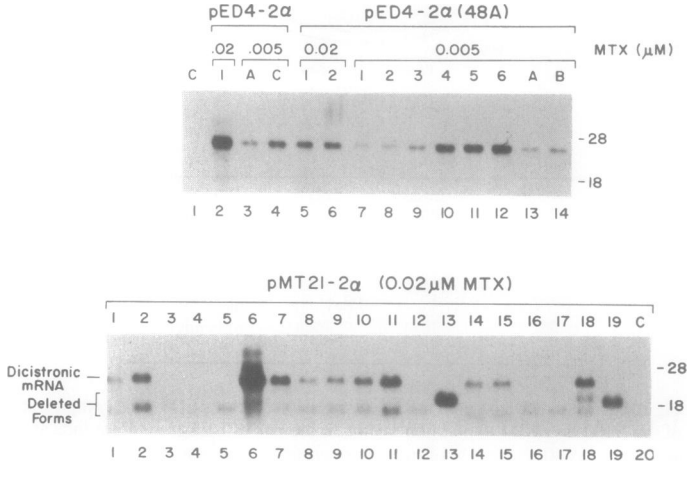

Figure 3. eIF-2α mRNA expression from pMT21 and pED4 vectors in stably transformed CHO cells. DHFR deficient GRA CHO cells were electroporated with linearized plasmid DNA from pED4-2α (top, lanes 2−4), pED4-2α(48A) (top, lanes 5−14), or pMT21-2α (bottom, lanes 1−19) as described in Table I and Materials and Methods. Individual clones (indicated by numbers above the lanes) or pools of clones (indicated by letters above lanes 3, 4, 13, and 14 on the top panel) were selected for resistance to the concentration of MTX indicated (0.02 or 0.005 μM). Total cell RNA was isolated and analyzed by Northern blot hybridization to a DHFR probe. The position of migration of 18S and 28S ribosomal RNA is shown. Lane 1 in the top panel and lane 20 in the bottom panel represent RNA from control untransfected GRA CHO cells.

detected. The higher molecular weight species corresponds to the correct sized dicistronic mRNA and also hybridizes to an eIF-2α probe (not shown). The multiple lower molecular weight mRNA species do not hybridize to an eIF-2α probe and likely represent mRNAs derived from a deleted or rearranged insert. The deletions which occur generate DHFR mRNAs which differ slightly in mobility between the different clones. In contrast, all clones as well as several pools of transformants derived from the pED4-2α vector express primarily a single mRNA of the expected size (Fig. 3, top). The same mRNA is also detected by hybridization to a eIF-2α probe (data not shown). Since a very minor lower molecular weight DHFR hybridizing RNA species may be detected in Fig 3A (lanes 2 and 5), this analysis cannot exclude the possibility that alternate monocistronic DHFR mRNAs are present at a low level as a result of alternate RNA processing or cryptic transcription initiation. By the mRNA analysis presented, there was no evidence of rearrangement or deletion of the vector in any of the transformants isolated using the pED4 vector upon selection for DHFR expression as compared to the pMT21 vector.

Amplification of pED4 Vectors in CHO Cells

Cells expressing eIF-2α from pED4 vectors were selected for amplification of the transfected DNA by sequential selection in increasing concentrations of MTX. At each level of methotrexate resistance, the expression of eIF-2α was characterized. The level of protein synthesis in the methotrexate resistant cells was evaluated by pulse labeling cells with ^{35}S-methionine and analysis of total protein by SDS-PAGE. eIF-2α represents the most abundant protein synthesized in the cells selected to 0.5 μM methotrexate (approximately 5% of the total protein synthesis) (Fig. 4, lane 4). Northern blot analysis of total cellular mRNA by hybridization to a DHFR probe demonstrated an 8-fold increase in mRNA expression for eIF-2α (Fig. 4, lanes 10–12) as MTX resistance selection was increased to 0.5 μM MTX. A single mRNA of the expected size was detected by hybridization to an eIF-2α probe.

Selection of pED4 Vectors in Cells that are DHFR Positive

The expression vector pED-mtx^r encodes a DHFR protein which has a leucine to arginine mutation at residue 22. As a consequence this protein is several hundred fold resistant to inhibition by MTX (10). The utility of this vector to transform and express a foreign gene was evaluated by insertion of a cDNA encoding a propeptide processing enzyme identified as furin or PACE (paired basic amino acid cleaving enzyme) (13,22). The protein encoded by this cDNA is a transmembrane subtilisin-like serine protease which can process precursor polypeptides (13). The resultant construct was electroporated into Rat-2 fibroblasts. Cells were selected by growth in 75, 150, and 300 nM MTX. The transformation frequency of the vector is shown in Table I.B. As the concentration of MTX increased from 75 to 300 nM there was a reduction in the background colonies. The frequency of transformation was 2×10^{-5} in 300 nM MTX. Clones isolated were characterized for the expression of PACE mRNA. Fig. 5 shows that the majority of these clones isolated express a single mRNA characteristic of the dicistronic mRNA. One out of seven lines has apparently rearranged the vector (clone 1).

DISCUSSION

The utilization of dicistronic mRNA expression vectors containing the desired gene in the 5′ position and a selectable marker gene in the 3′ position has proven a successful approach to derive cells that express high levels of the gene in the 5′ position (4,6). Because the translation of the selectable gene in the 3′ position is very inefficient, cells require higher levels of dicistronic mRNA to survive selection and the protein expressed from the coding region in the 5′ position is translated preferentially. However,

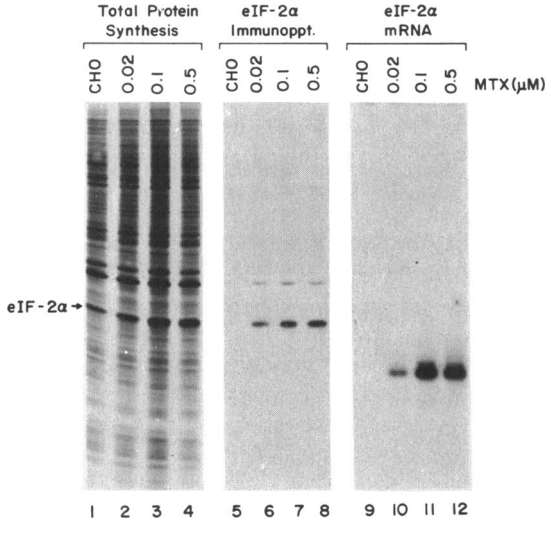

Figure 4. Amplification of eIF-2α expression by selection in increasing concentrations of MTX. CHO cells were electroporated with pED4-2α as in Table I and a clone isolated and propagated sequentially in 0.02, 0.1, and 0.5 μM MTX. At each concentration cells were pulse labeleed with ^{35}S-methionine and extracts prepared for SDS-PAGE analysis before (lanes 1–4) and after (lanes 5–8) immunoprecipitation with an anti-eIF-2α specific antibody as described in Materials and Methods. In parallel, total RNA was prepared for analysis by hybridization to a DHFR probe (lanes 9–12).

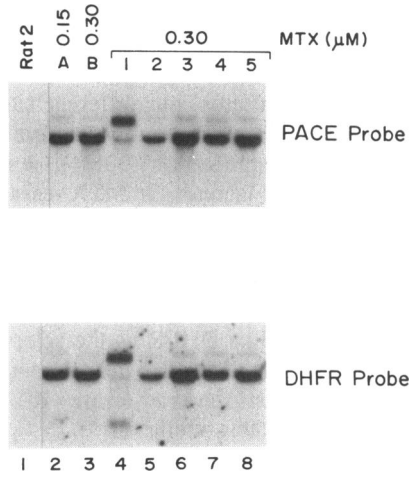

Figure 5. PACE mRNA expression from pED-mtx^r-PACE in Rat-2 cells. Rat-2 fibroblasts were electroporated with pED-mtx^r-PACE and selected for resistance to 0.15 or 0.30 μM MTX as described in Table I. Total cell RNA was prepared from two pools of transformants (A and B in lanes 2,3) and 5 clones (1–5 in lanes 4–8) for analysis by Northern blot hybridization to a DHFR probe (bottom) or a PACE probe (top) as described in Materials and Methods. Lane 1 represents RNA from original Rat-2 fibroblasts.

the general utility of this approach has been limited for several reasons. First, since the frequency of initiation at the 3′ selectable gene is approximately 1% of the frequency of initiation at the 5′ coding region (4,5), in many cases it is not possible to produce adequate levels of the selection marker product to survive selection. Second, insertion of sequences upstream from the selection marker gene may dramatically alter the efficiency of initiation at the downstream AUG codon for the selection marker. The distance between the translation termination codon of the upstream coding region and the next AUG should optimally be 100 bases and should be void of AUG codons (5). Finally, our experience with methotrexate selection for amplification of DHFR expression in dicistronic mRNA vectors has demonstrated that frequently deletion of the upstream sequences occurs, presumably in order to allow more efficient translation of DHFR. An example of the frequency by which rearrangements likely occur is presented in Fig. 3. Using previously described dicistronic mRNA expression vectors it is difficult to know whether these deletions occur as result for selection for increased DHFR expression or as a result of selection against high level expression of the protein encoded within the 5′ end of the mRNA.

The series of pED vectors described in this report overcome the limitations of previously used dicistronic mRNA expression vectors. Incorportion of an internal ribosomal entry site to promote more efficient internal translation initiation at the DHFR selectable marker gene has increased the efficiency of transformation and amplification of the pED4 vector in CHO cells that are DHFR deficient. In addition, the pED-mtx^r vector which harbors a leucine to arginine mutation at residue 22 of DHFR and encodes an enzyme that is several hundred fold resistant to MTX efficiently transforms cells that are not DHFR deficient. The pED vectors have demonstrated success in efficiently, rapidly, and reproducibly obtaining stable cell lines which express high levels of protein expressed from heterologous genes. These vectors should prove advantageous over other stable selection expression vectors for several reasons. First, since DHFR is more efficiently translated, selection for deletion of sequences inserted into the 5′ position is diminished. If deletions do occur with a particular insert, it may be an indication that those sequences or the translated protein is detrimental to cell viability. Second, since these vectors yield greater transformation frequencies, it is possible to pool large numbers of transformants that have integrated the vector into different sites within the genome. Since different integration sites have dramatically different potential for amplification (23,24), it is possible to rapidly select by sequential increases in MTX resistance to obtain the subpopulation of cells that rapidly amplify the gene to high copy number. Our experience is that selection of pools of transformants to MTX resistance can yield high amplification of expression within 8 weeks using these vectors. Although the results presented here utilize DHFR as the selectable marker, we have constructed similar EMC virus leader containing vectors which encode neomycin phosphotransferase (16) which can be used as a dominant selectable marker.

The vectors described in this report should be of general utility for the stable expression of foreign genes in mammalian cells. As shown here these vectors are readily amplifiable to obtain high level expression of foreign genes. Introduction of the EMC virus leader sequence does not interfere with the translation of ORFs inserted upstream of the EMC virus leader upon transient transfection of COS-1 cells. Consequently, these vectors may by used to efficiently obtain expression by transient DNA transfection of COS-1 cells as as well as stable transfection and subsequent amplification in CHO cells. Since the inserted sequences do not influence the expression of the selectable marker, these vectors should also be useful for the construction of cDNA expression libraries in stably transformed cell lines, similar to those described using Epstein Barr virus derived vectors (25). These vectors should also prove useful for expression of anti-sense RNA by inserting sequences upstream from the EMC leader. Finally, these vectors will improve the ability to select for homologous recombination in sequence replacement-type vectors by inserting DNA sequences to be targeted upstream on the EMC leader sequence. This latter approach has been used to switch a mouse heavy chain mu immunoglobumin gene to a human gamma 1 heavy chain (16).

REFERENCES

1. Kozak, M. (1978) Cell *15*: 1109–1123.
2. Kozak, M. (1989) J. Cell Biol. *108*: 229–241.
3. Peabody, D.S., and Berg, P. (1986) Mol. Cell. Biol. *6*: 2695–2703.
4. Kaufman, R.J., Murtha, P., and Davies, M.V. (1987) EMBO J. *6*: 187–193.
5. Kozak, M. (1987) Mol. Cell. Biol. *7*: 3438–3445.
6. Balland, Faure, T., Carvallo, D., Cordier, P., Ulrich, P., Fournet, B., De La Salle, H., and Lecocq, J-P. (1988) Eur. J. Biochem. *172*: 565–572.
7. Pelletier, J., and Sonenberg, H. (1988) Nature (London) *334*: 320–325.
8. Jang, S.K., Davies, M.V., Kaufman, R. J., and Wimmer, E. (1989) J. Virol. *63*: 1651–1660.
9. Jang, S.K., and Wimmer, E. (1990) Genes and Dev. *4*: 1560–1572.
10. Simonsen, C.C., and Levinson, A.D. (1983) Proc. Natl. Acad. Sci. USA *80*: 2495–2499.
11. Kaufman, R.J., Davies, M.V., Pathak, V.K., and Hershey, J.W.B. (1989) Mol. Cell. Biol. *9*: 946–958.
12. Wise, R.J., Barr, P.J., Wong, P.A., Kiefer, M.C., Brake, A.J., and Kaufman, R.J. (1990) Proc. Natl. Acad. Sci. USA *87*: 9378–9382.
13. Gluzman, Y. (1981) Cell *23*: 175–182.
14. Kaufman, R.J. (1990) Technique *2*: 221–236.
15. Israel, D.I., and Kaufman, R.J. (1989) Nucl. Acids Res. *12*: 4589–4604.
16. Wood, C.R., Dorner, A.J., Morris, G.E., Alderman, E.M., Wilson, D., O'Hara, R.M., and Kaufman, R.J. (1990) J. Immunol. *145*: 3011–3016.
17. Laemmli, U.K. (1970) Nature *227*: 680–685.
18. Chirgwin, J.M., Przybyla, A.E., R.J. MacDonald, and Rutter, W.J. (1979) Biochemistry *18*: 5294–5299.
19. Thomas, P. (1980) Proc. Natl. Acad. Sci. USA *77*: 5201–5205.
20. Derman, E., Krauter, K., Walling, L., Weinberg, C., Ray, M., and Darnell, Jr.,J. R. (1981) Cell *23*: 731–739.
21. Kaminski, A., Howell, M.T., and Jackson, R.J. (1990) EMBO J. *9*: 3753–3759.
22. Fuller, R.S., Brake, A.J., and Thorner, J. (1989) Science *246*: 482–486.
23. Wahl, G.M., Vincent, B.R.S., and DeRose, M.L. (1984) Nature 307: 1051–1057.
24. Kaufman, R.J., and Sharp, P.A. (1982) J. Mol. Biol. *159*: 601–621.
25. Margolskee, P., Kavathas, and Berg, P. (1988) Mol. Cell. Biol. *8*: 2837–2847.
26. Kaufman, R.J. and Sharp, P.A. (1983) Mol. Cell. Biol. *2*: 1302–1319.