Volume 219, number 1, 181–188 FEB 04898 July 1987

Expression of human pancreatic polypeptide precursors from a dicistronic mRNA in mammalian cells

Esper Boel, Kathleen L. Berkner°, Bjørn A. Nexø and Thue W. Schwartz*

*Novo Research Institute, Laboratory of Genetics, DK-2880 Bagsværd-Copenhagen, *Laboratory of Molecular Endocrinology, University Department of Clinical Chemistry, Rigshospitalet 6321, DK-2100 Copenhagen, Denmark* and *°ZymoGenetics, Inc., 2121 North 35th Street, Seattle, WA 98103, USA*

Received 6 May 1987

The ability of eukaryotic ribosomes to reinitiate translation at downstream AUG codons on polycistronic mRNAs was used to select transfected CHO clones that secreted a precursor to the human pancreatic polypeptide (PP). In the in vitro constructed transcription unit, a viral promoter directed the synthesis of a dicistronic mRNA. The PP cDNA was placed in the 5′-part of this transcript, while a DHFR cDNA was placed 3′ to the PP. This dicistronic expression unit was transfected into CHO cells, and methotrexate-resistant colonies were selected. RNA-blots verified that the PP precursor and DHFR were expressed from the same dicistronic mRNA. The CHO cells synthesized the hormone precursor and secreted it through the constitutive secretory pathway.

Dicistronic transcript; Cell transfection; Open reading frame; Adenoviral promoter

1. INTRODUCTION

A growing number of recent investigations suggests that eukaryotic ribosomes have the capacity to reinitiate translation at a downstream (internal) AUG, once they have terminated translation of an upstream open reading frame [1–3]. The hypothesis for translation initiation in eukaryotes [4–6] states that the 40 S ribosomal subunit first binds at the 5′-capped end of the mRNA. A scanning mechanism is then suggested, in which the 40 S subunit migrates [7] through the 5′-untranslated region until an AUG is reached. In approx. 94% of 221 inspected eukaryotic mRNAs [8] the first AUG encountered by the migrating 40 S subunit serves as the initiator. Eukaryotic ribosomes have a strong preference [5] for initiating translation at AUGs located in the consensus sequence ACCAUGG. When the 40 S subunit reaches an AUG in this consensus, the entire 80 S ribosome will assemble and initiate translation. In the few per cent of eukaryotic mRNAs where translation of the encoded protein does not start at the 5′-proximal AUG, this is most often correlated with the presence of an unfavourable sequence surrounding this AUG, e.g. if the −3 position contains a pyrimidine. In these cases the 40 S subunit will continue scanning until it reaches internal AUGs further downstream.

In experimentally conceived mRNAs [1–3,9,10] it has been demonstrated that initiation at a downstream AUG can be severely impaired if this AUG is preceded by a favourable upstream AUG. If translation from the 5′-proximal AUG continues beyond the downstream AUG, translation from the downstream AUG can be totally abolished. However, if a terminator codon stops the translation initiated at the 5′-proximal AUG before it reaches the downstream AUG, the latter codon can now serve as a translational start site.

Correspondence address: E. Boel, Novo Research Institute, Laboratory of Genetics, DK-2880 Bagsværd-Copenhagen, Denmark

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

Thus termination-reinitiation is a feature of the translational process, which is not confined to the well-known polycistronic systems in prokaryotes. This new realization opens up a reinterpretation of data on gene expression from both cellular and viral polycistronic mRNAs, where potential internal AUG initiation codons are found.

Termination-reinitiation may also find advantageous applicability in expression systems designed for eukaryotic cells. In transfection experiments with mammalian cells a selectable genetic marker is most often introduced, together with the expression system under investigation, into the target cells on either a separate plasmid or on the same plasmid as the investigated heterologous expression system, but under control of a separate promoter. In the present report we present a new strategy for expression vector design. A dicistronic transcription unit, which directs synthesis of the precursor of human pancreatic polypeptide (PP) [11] as well as the synthesis of a selectable genetic marker (the mouse DHFR), can be expressed in mammalian cells.

2. MATERIALS AND METHODS

2.1. *Plasmid constructions*

The plasmids used in this investigation, pPP, pDHFR^r-I and pPP-DHFR^r, were constructed by the use of standard techniques [12], and their structural elements are described in section 3.1.

2.2. *Cell culture and DNA transfection*

Chinese hamster ovary (CHO) cells and clones derived from this cell line were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum at 5% CO_2. Plasmid DNA was transfected into cells using calcium phosphate precipitation [13] and 48 h post-transfection cells were split at an approximate 1:8 ratio into new media supplemented with methotrexate (0.3 μM Mtx). In a typical transfection experiment using 50% confluent cells in a 9 cm culture dish, approx. 10 μg of plasmid was used. pPP-DHFR^r was used alone and pPP and pDHFR^r-I were used together in a 10:1 molar ratio. Clones were picked after 10 days and propagated in media with 0.3 μM Mtx.

2.3. *DNA and RNA extraction and analysis*

DNA was extracted from cell lines and 10 μg was digested with different restriction enzymes before electrophoresis on 1% agarose gels. After Southern blot transfer [14] the filters were hybridized to a DHFR specific probe labelled by nick translation [15]. Total cellular RNA was extracted [16] from CHO cell lines and was analysed with ^{32}P-nick-translated [15] DHFR- and PP-specific DNA probes after electrophoresis in formaldehyde containing agarose gels [12] as described in [17].

2.4. *Radioimmunoassay of PP*

PP was measured by radioimmunoassay as described [18] in medium from cell incubations directly and in cell extracts. Peptides were extracted from pellets of approx. 10^7 cells by boiling for 2 min in 2 ml of 50% acetic acid. Total protein was determined in the acetic acid extract directly by the Bio-Rad, Bradford method [19].

2.5. *Peptide characterization*

^{35}S-Methionine-labelled peptides were charac-

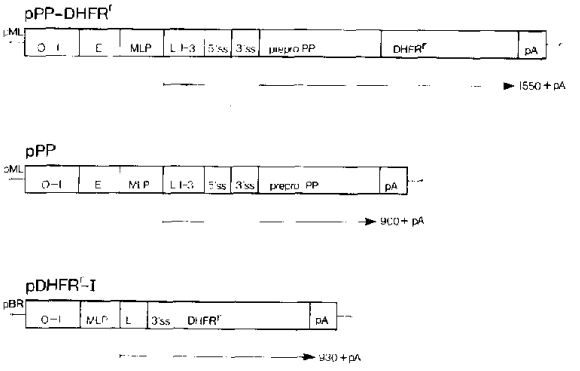

Fig.1. The transcription units used in three mammalian expression plasmids are shown as a number of open boxes each representing a functional unit. Each transcription unit was constructed on either pBR322 or pML [34]. '0-1' designates 0 to 1 map unit of adenovirus 5; 'E' is the SV40 early enhancer; 'MLP' is the adenovirus 2 (Ad2) major late promoter; 'L 1–3' is the tripartite leader from Ad2; '5'ss' is a 5'-splice site; '3'ss' is a 3'-splice site; 'pA' is the polyadenylation signal from the late transcriptional unit of SV40 in pPP-DHFR^r and in pPP. In pDHFR^r-I the early polyadenylation signal is used. pDHFR^r-I lacks the SV40 enhancer and contains only the first Ad2 leader. The expected spliced transcript is shown below each DNA construction.

terized by SDS-gel electrophoresis [20] after immunoprecipitation. Cells in a 90% confluent 160 cm^2 flask were incubated with ^{35}S-methionine for 4 h in medium lacking methionine. Radiolabelled peptides were extracted from the medium by Sep-Pak purification. Immunoprecipitation was performed in 50 mmol/l Tris, pH 7.4, containing 0.2 mol/l NaCl, 6 mmol/l EDTA, 5% (v/v) aprotinin solution, and 1.25% (v/v) Triton X-100. Two different PP antisera, 2 μl for each sample, were used for immunoprecipitation: Ab 146-7, a gift from Dr R.E. Chance (this antiserum is known to precipitate the PP precursor [21]), and Ab 205, a gift from Dr M.M.T. O'Hare [22]. After 20 h of incubation the precipitation was performed by protein A-Sepharose and the radiolabelled peptides were extracted by boiling in the electrophoresis sample mixture. The specificity of the immunoprecipitation was tested by parallel experiments where 50 μg of bovine PP was added. Pulse-chase experiments were performed on cells in 2.5 cm petri dishes, incubated for 15 min with 0.25 mCi of ^{35}S-methionine each. Radiolabelled peptides were isolated from the medium, as described above, after different chase periods in medium containing unlabelled methionine.

3. RESULTS

3.1. *PreproPP and DHFR^r expression plasmids*

The three adenovirus promoter based expression plasmids used in the present study are shown in fig.1. In pDHFR^r-I [23], the mouse DHFR cDNA has been mutated by oligonucleotide-directed site specific mutagenesis to change the Leu-22 to an Arg. This mutation gives DHFR the Mtx-resistant phenotype [24]. The human preproPP cDNA from pHPP165 [11] was isolated as a 570 bp *Pst*I fragment. pPP was generated by adding *Bam*HI linkers to the preproPP encoding fragment which was then ligated into the *Bam*HI cloning site of pD5. The structure of pD5 is similar to pD11 [25]. In pPP-DHFR^r the promoter directs the synthesis of a dicistronic mRNA that encodes both preproPP and DHFR^r. The *Fnu*4H I site just 5′ to the initiator codon of DHFR^r [24] in pDHFR^r-I was changed to a *Bam*HI site by linker addition, deleting most of the 5′-nontranslated sequences from DHFR^r cDNA. Cells transfected with this dicistronic construct can only grow in Mtx, if DHFR^r is expressed from a transcribed dicistronic mRNA, which also directs the synthesis of preproPP, or if some gene rearrangement has taken place that locates the open reading frame of DHFR^r at a more 5′-position on the mRNA.

3.2. *Isolation of HPP-producing cell lines*

CHO cells were transfected by the calcium phosphate technique [13] with pPP-DHFR^r or with

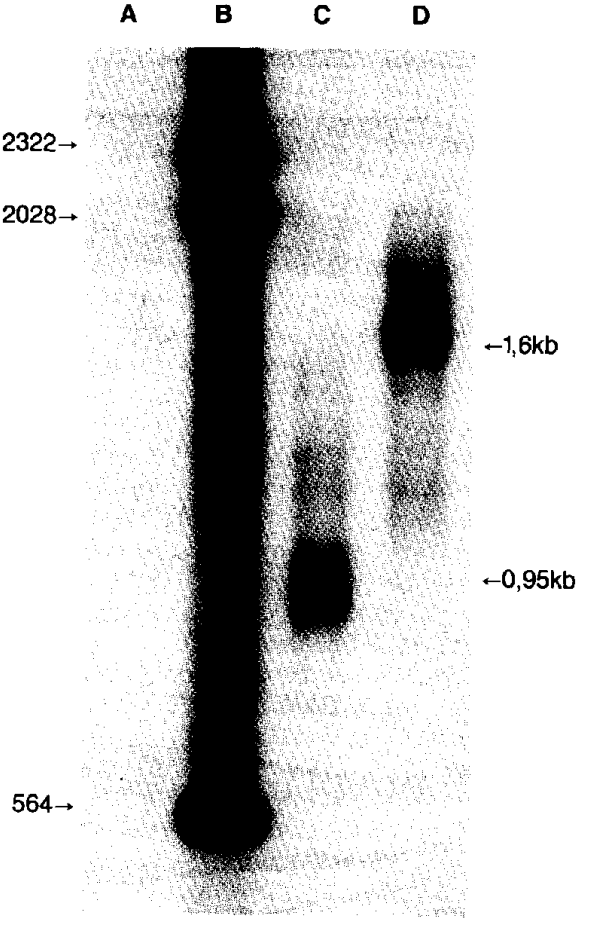

Fig.2. Autoradiography from blot analysis of CHO cell RNA. 40 μg of total cellular RNA was applied to each lane on a 1.2% formaldehyde agarose gel. After electrophoresis the RNA was blotted onto nitrocellulose and hybridized with ^{32}P-nick-translated DHFR^r DNA. Lanes: A, RNA from nontransfected cells; B, bacteriophage DNA digested with *Hin*dIII and labelled with ^{32}P; C, RNA from cell line CHO-571 (cotransfected with pDHFR^r-I and pPP); D, RNA from cell line CHO-728 (transfected with pPP-DHFR^r).

pPP together with pDHFR^r-I. No colonies were detected in mock transfected cells grown in 0.3 μM Mtx. After 10 days clones could be isolated from cells that had been cotransfected with pPP and pDHFR^r-I as well as from cells transfected with the dicistronic construct pPP-DHFR^r. pPP-DHFR^r gave approximately 20-fold lower cloning efficiency than pDHFR^r-I when an equivalent amount of plasmid DNA was used in the transfection experiments. All colonies generated by this dicistronic gene secreted measurable amounts of PP immunoreactive peptides. DNA, RNA and immunoreactive peptides were analysed in a representative clone from the cotransfection experiment (CHO-571) and in one from the transfection with pPP-DHFR^r (CHO-728).

3.3. *Analysis of DNA and RNA from transfected cell lines*

DNA extracted from clones CHO-571 and CHO-728 was digested with a number of restriction enzymes and analysed by Southern blotting [14] using ^{32}P-nick-translated DHFR fragments as hybridization probes. We could detect no rearrangements in the DNA from CHO-571 nor in the DNA from CHO-728 as compared to the structure of the introduced expression units (not shown). It is possible, however, that splicing events could have shifted the DHFR^r open reading frame to a more 5′-location on a rearranged mRNA molecule to generate a monocistronic unit, and we therefore carried out Northern blot analysis on extracted RNA from the two cell lines, with both DHFR^r and PP specific probes. In fig.2 the autoradiography from a Northern blot on RNA from untransfected CHO cells and from the cell lines CHO-571 and CHO-728 is shown. In this analysis the RNAs were hybridized to a ^{32}P-nick-translated DHFR^r restriction fragment. The 0.95 kb mRNA in lane C corresponds to the DHFR^r specific transcript synthesized from the pDHFR^r-I gene (fig.1), which was used in a cotransfection with pPP to generate the cell line CHO-571. The CHO-728 clone was obtained from a transfection with the dicistronic pPP-DHFR^r alone, and DHFR-specific mRNA from this cell line (lane D in fig.2) was 1.6 kb long, which is in good agreement with the expected length of a dicistronic transcript from the PP-DHFR^r gene (fig.1). In fig.3 the same RNAs were analysed for PP-specific transcripts. In lane C a

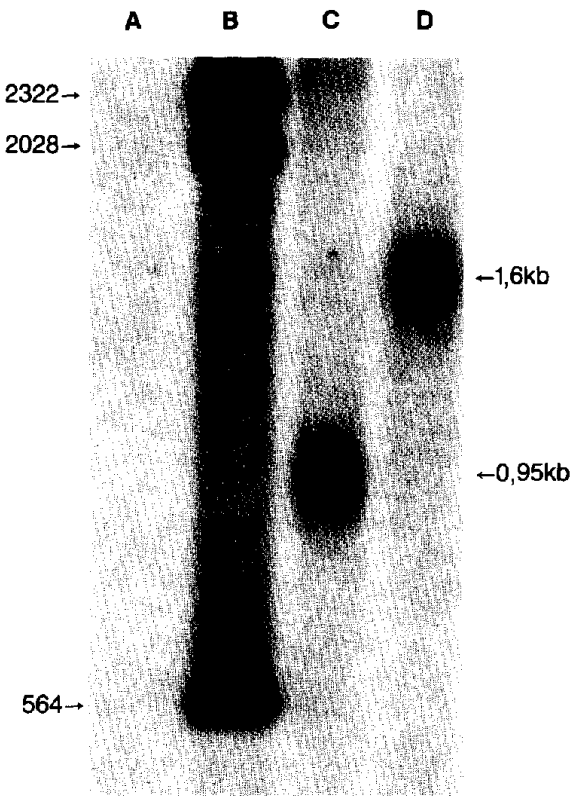

Fig.3. Autoradiography from blot analysis of CHO cell RNA. Electrophoresis and RNAs applied to the different lanes were as in fig.2. The blot was hybridized with ^{32}P-nick-translated PP DNA.

0.95 kb mRNA is identified, and this size fits well with the expected 0.9 kb transcript from the monocistronic PP construct (fig.1). In lane D a 1.6 kb PP specific mRNA is observed, and the size of this RNA species corresponds precisely to the DHFR^r specific transcript detected in the same cell line (fig.2, lane D). It can be concluded that CHO-728 expresses DHFR and PP from the same dicistronic 1.6 kb mRNA.

3.4. *Peptide characterization*

In table 1 the amounts of PP stored in and secreted from transfected cells are shown. Labelling experiments with ^{35}S-methionine showed that the PP produced in the CHO cells as judged by SDS-gel electrophoresis corresponded in size (9.5 kDa) to the PP precursor previously characterized (fig.4) [11]. During prolonged in-

Table 1

Pancreatic polypeptide synthesis and secretion from Chinese hamster ovary (CHO) cells transfected with a dicistronic vector, pPP-DHFR^r, containing both the PP and the DHFR coding sequences, or cells cotransfected with separate vectors, pPP and pDHFR^r-I

	Transfection type	Content PP molecules per cell	Secretion PP molecules per cell per 24 h
CHO-wild type	–	0	0
CHO-571[a]	cotransfect.	3.6×10^5	$2.6–3.0 \times 10^{6b}$
CHO-727	dicistronic	N.D.	5.5×10^5
CHO-728[a]	dicistronic	3.8×10^5	$0.9–1.1 \times 10^{6b}$
CHO-731	dicistronic	N.D.	1.2×10^5

[a] Cell lines in which the DNA and RNA were studied by Southern and Northern analysis

[b] Secretion studied on two occasions six months apart

N.D., not determined

cubation with labelled amino acid, a somewhat larger (15 kDa), labelled peptide is also detected (figs 4 and 5); it has not been clarified in the present study whether this reflects a modification of the PP precursor occurring in the medium or in the cells. It should be noted that a glycosylation site is found in the precursor at amino acid residues Asn-40–Ala–Thr-42 [11]. During pulse-chase experiments the precursor was secreted rapidly, within 30–45 min, into the medium (fig.5). This rapid secretion of unprocessed precursor is in accordance with the fact that the CHO cell is not a neuroendocrine cell type, and therefore is not capable of storing peptide precursors and secreting them through the so-called regulated pathway [26].

4. DISCUSSION

We have demonstrated that dicistronic expression units can be used to select transformed cell lines. If the selectable genetic marker, in this study an Mtx-resistant mutant of DHFR, is incorporated as a second open reading frame on the dicistronic mRNA, this genetic marker is apparently expressed at a sufficiently high efficiency to confer an Mtx-resistant phenotype on transformed cells. If a cDNA is built into the dicistronic construction as a 5′-proximal open reading frame, clones expressing the selectable marker will express the cDNA. We have verified by Northern blotting that both

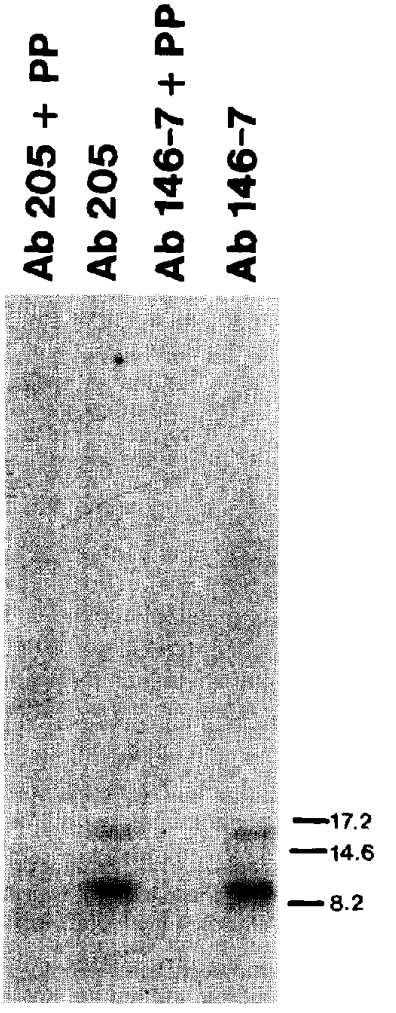

Fig.4. Characterization of immunoreactive PP peptides by immunoprecipitation and SDS-gel electrophoresis. Media from cells, clone CHO-728, incubated with ^{35}S-methionine, were extracted by Sep-Pak and radiolabelled peptides were immunoprecipitated by two PP-antisera, 146-7 and 205, with and without the addition of an excess of unlabelled PP. The precipitation was performed with an excess protein A-Sepharose. Radiolabelled peptides were eluted from the beads with SDS sample mixture before electrophoresis in 16% SDS-acrylamide slab gels. The precursor for PP (the prohormone) consists of 66 amino acids and migrates with an apparent molecular mass of 9500 Da in SDS gels [11]. The position of size markers is shown.

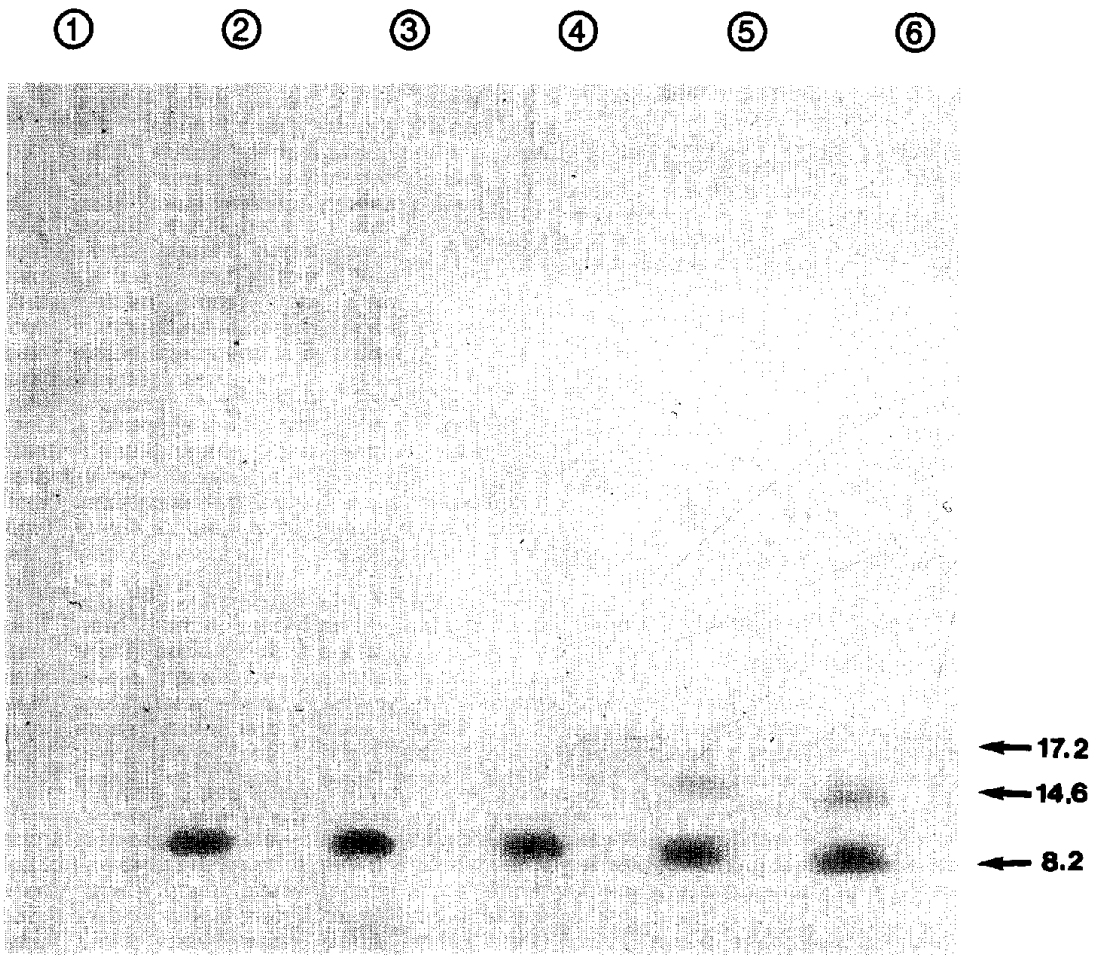

Fig.5. SDS-gel electrophoresis of immunoprecipitable PP peptides in media from CHO-571 cells during a pulse-chase experiment. Cells were labelled with ^{35}S-methionine for a pulse period of 30 min (lane 1) followed by incubation with excess amount of unlabelled methionine for 15 min (lane 2), 30 min (lane 3), 60 min (lane 4), 120 min (lane 5), and for 210 min (lane 6).

PP and DHFR are encoded on the same 1.6 kb mRNA transcript. Peabody et al. [2,3] confirmed that eukaryotic ribosomes can reinitiate translation by analysis of translated proteins synthesized from constructed dicistronic transcripts, on which two open reading frames were positioned in a number of different arrangements with respect to each other. This possibility had previously been suggested, e.g. by deletion mutations affecting translation in plants of the open reading frames VII and I of the cauliflower mosaic virus [27], and by analysis of the effects of upstream AUG and stop codons on the translation of the hepatitis B virus surface antigen [10] and the *E. coli* galactokinase gene (galK) [1] in mammalian cells.

This ability of eukaryotic ribosomes to reinitiate translation offers the possibility of designing expression units with a close functional and structural connection between the DNAs of the investigated protein and the selectable marker. The fact that expression of a heterologous gene product is closely coupled to the expression of a selectable genetic marker transcribed from the same promoter might offer advantages in long term clone stability and in the amplification potential of the integrated gene. During preparation of our

manuscript a similar study on the use of polycistronic vectors in the direct selection of transfected cell lines was published by Kaufman et al. [28]. They used the wild type mouse DHFR as a selectable marker in CHO DUKX-B11 cells [29], and demonstrated expression of adenosine deaminase (ADA) and granulocyte macrophage colony stimulating factor (GM-CSF) from dicistronic constructs. In COS cells [30] they also demonstrated expression from a tricistronic transcript which included the three above-mentioned cDNAs. Human plasminogen activator cDNA (tPA) has also been expressed [31] from a dicistronic construct in CHO DUKX-B11 cells, and we have obtained expression of human clotting factor IX (FIX), *E. coli* chloramphenicol acetyltransferase (CAT) and mouse DHFR from di- and tricistronic vectors transfected into BHK cells (Berkner et al., in preparation). The neomycin gene (neo) from Tn5 [32] can be used as well to obtain stable transformants from polycistronic transcripts. Recently we used G418 selection together with both di- and tricistronic constructs encoding human neuropeptide Y (NPY) [33], PP and *neo* to obtain stable transformants in the neuroendocrine cell line AtT-20 (Wulff et al., in preparation). The growing number of different cDNAs of variable nucleotide lengths, which have been incorporated into functional eukaryotic polycistronic transcription units, argues that this concept of expression vector design might have general applicability.

Our original expectations to the performance of the dicistronic constructs were that by weakening the translation of the selectable genetic marker by incorporating it as a second open reading frame on a dicistronic mRNA, we would enforce an extra hard selection on transfected cells, thereby generating ones that transcribed a relatively large amount of dicistronic mRNA. This could be obtained by, e.g., chromosomal integration in transcriptionally very active regions of the genome, and this would be reflected in a reproducibly higher level of product synthesized from the 5′-proximal open reading frame. The significantly fewer colonies generated with polycistronic expression units (approx. 20-fold in our study) as compared to the monocistronic type shows that expression of the selectable marker is impaired. However, we have consistently found no higher level of expression from the 5′-proximal open reading frame in the di- or tricistronic constructs. Apparently there is still much performance evaluation to be done on this new type of polycistronic eukaryotic transcription units, before their full potential can be evaluated. We have initiated such work, which will involve, for instance, changes in the intercistronic regions, mutations of the sequences around initiating AUG codons, gene amplification and long term stability studies. The results from such studies will be reported elsewhere (Berkner et al., in preparation).

ACKNOWLEDGEMENTS

The technical assistance of Ms Iben Hjort and Ms Lone Bredo Pedersen is gratefully acknowledged. The manuscript was typed with the usual expertise by Ms Marit Petersen. E.B. would like to thank Dr Niels P. Fiil, Vice President, Molecular Biology, Novo Research Institute, and the mammalian expression group at ZymoGenetics, Inc., Seattle for the opportunity to work at ZymoGenetics in the winter of 1984/1985, during which period the work on eukaryotic polycistronic gene units was initiated.

REFERENCES

[1] Johansen, H., Reff, M., Schumperli, D. and Rosenberg, M. (1984) in: Gene Expression, Alfred Benzon Symposium 19 (Clark, B.F.C. and Petersen, H.U. eds) pp.413–429, Munksgård, Copenhagen.
[2] Peabody, D.S. and Berg, P. (1986) Mol. Cell Biol. 6, 2695–2703.
[3] Peabody, D.S., Subramani, S. and Berg, P. (1986) Cell Biol. 6, 2704–2711.
[4] Kozak, M. (1978) Cell 15, 1109–1123.
[5] Kozak, M. (1986) Cell 44, 283–292.
[6] Kozak, M. (1980) Cell 22, 7–8.
[7] Kozak, M. and Shatkin, A.J. (1978) J. Biol. Chem. 253, 6568–6577.
[8] Kozak, M. (1984) Nucleic Acids Res. 12, 857–872.
[9] Kozak, M. (1984) Nucleic Acids Res. 12, 3873–3893.
[10] Liu, C.-C., Simonsen, C.C. and Levinson, A.D. (1984) Nature 309, 82–85.
[11] Boel, E., Schwartz, T.W., Norris, K.E. and Fiil, N.P. (1984) EMBO J. 3, 909–912.

[12] Maniatis, T., Fritsch, F.F. and Sambrook, J. (1982) A Cloning Manual, Cold Spring Harbor Laboratory Press, New York.
[13] Graham, F.L. and Van der Eb, A.J. (1973) Virology 52, 456–467.
[14] Southern, E.M. (1975) J. Mol. Biol. 98, 503–518.
[15] Rigby, P.W.J., Dickmann, M., Thodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237–251.
[16] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W. (1979) Biochemistry 18, 5294–5299.
[17] Boel, E., Bech, A.-M., Randrup, K., Dræger, B., Fiil, N.P. and Foltmann, B. (1987) Proteins, in press.
[18] Schwartz, T.W., Fahrenkrug, J., Holst, J.J., Jensen, S.L., Nielsen, O.V., Rehfeld, J.F., Muckadell, O.B.S. and Stadtl, F. (1976) J. Clin. Invest. 61, 781–789.
[19] Bradford, M. (1976) Anal. Biochem. 72, 248–256.
[20] Laemmli, U.K. (1970) Nature 227, 680–685.
[21] Schwartz, T.W., Gingerich, R.L. and Tager, H.S. (1980) J. Biol. Chem. 255, 11494–11498.
[22] O'Hare, M.M.T., Daly, J.G. and Buchanan, K.D. (1983) Clin. Chem. 29, 1923–1927.
[23] Berkner, K.L. and Sharp, P.A. (1984) Nucleic Acids Res. 12, 1925–1941.
[24] Simonsen, C.C. and Levinson, A.D. (1983) Proc. Natl. Acad. Sci. USA 80, 2495–2499.
[25] Powell, J.S., Berkner, K.L., Lebo, R.V. and Adamson, J.W. (1986) Proc. Natl. Acad. Sci. USA 83, 6465–6469.
[26] Kelly, R.B. (1985) Science 230, 25–32.
[27] Dixon, L.K. and Hohn, T. (1984) EMBO J. 3, 2731–2736.
[28] Kaufman, R.J., Murtha, P. and Davies, M.V. (1987) EMBO J. 6, 187–193.
[29] Urlaub, G. and Chasin, L.A. (1980) Proc. Natl. Acad. Sci. USA 77, 4216–4220.
[30] Gluzman, Y. (1981) Cell 23, 175–182.
[31] Kaufman, R.J., Wasley, L.C., Spiliotes, A.J., Gossels, S.D., Latt, S.A., Larsen, G.R. and Kay, R.M. (1985) Mol. Cell Biol. 5, 1750–1759.
[32] Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. and Schaller, H. (1982) Gene 19, 327–336.
[33] Minth, C.D., Bloom, S.R., Polak, J.M. and Dixon, J.E. (1984) Proc. Natl. Acad. Sci. USA 81, 4577–4581.
[34] Lusky, M. and Botchan, M. (1981) Nature 293, 79–81.