Vector Design for Optimal Protein Expression

BioTechniques 31:572-282 (September 2001)

Alike W. van der Velden, Harry O. Voorma, and Adri A.M. Thomas
Utrecht University, Utrecht, The Netherlands

ABSTRACT

Many DNA constructs are generated for protein expression studies. Translational properties and mRNA stability are crucial aspects that have to be accounted for during DNA construction. An optimized vector for protein overexpression studies is described considering elements in the mature mRNA that influence translatability and stability. Recommendations regarding vector construction for Xenopus laevis *embryo injection are provided, based on literature and experimental data. The 5′ untranslated region (5′UTR) should be non-regulated, short, unstructured, and without AUG codons. The sequence around the start codon should match the initiation context of the species studied (ACCAUGG, for vertebrates), and the open reading frame should be cloned with its own stop codon, followed by a G or A residue. Furthermore, the 3′UTR should be non-regulated, and a strong polyadenylation signal must be included in DNA vectors. In RNA template vectors, the presence of a poly(A) or AC tail is essential for stability, as well as for translation efficiency in mRNA injection experiments. These aspects result in high-level expression of exactly the desired protein. Easily obtainable examples of the sequences [5′UTR, 3′UTR, and poly(A) signal] are suggested.*

INTRODUCTION

In molecular biology research, many DNA constructs are generated for protein expression studies. Protein overexpression is used to elucidate the biological function of the protein of interest and the process it is functioning in. Two expression systems are often used: transient or stable transfection of tissue culture cells, or injection in a one-cell stage embryo.

Transfection of tissue culture cells requires a DNA construct with the gene of interest cloned under the control of a constitutive or inducible promoter. This can be the gene's own promoter, but strong viral promoters, like the rous sarcoma virus (RSV) or cytomegalovirus (CMV) promoters, are often used. Two types of genetic information can be injected into embryos, mRNA or DNA. First, injection of synthetic mRNA encoding the protein of interest requires the construction of a DNA construct with a promoter active in in vitro transcription, such as the T_7, SP6, or T3 phage promoters. The advantage of injecting RNA is that protein expression starts immediately after injection in the vast majority of embryonic cells. Second, DNA constructs containing the gene of interest can be injected; this requires the use of a promoter active in the embryo. When protein expression is required in specific embryonic cells or tissues, a tissue-specific promoter can be used. In contrast, when the protein has to be overexpressed in the whole embryo, a general promoter—such as actin, histone, or a viral promoter—can be used. The advantages of DNA injection are that protein expression can be targeted by selection of the promoter, and there is no need for RNase-free work. The disadvantage is that protein expression starts after transcription initiation in the embryo (in the often used *Xenopus laevis* embryo after nine cleavages, about 10 h). Furthermore, injected DNA does not segregate as well as mRNA, resulting in a very mosaic expression pattern (44).

Besides the choice of the promoter and the gene of interest, there are many more important aspects that need to be considered in vector design. Often, the gene is cloned into a commercially available or laboratory-made vector, without considering aspects such as translatability and stability of the resulting mRNA.

First, we will introduce translation initiation, the first step in protein synthesis, as this step largely determines the efficiency of mRNA translation (23). The mature mRNA consists from 5′→3′ of a 7-methyl guanosine cap-structure, a 5′ untranslated region (5′UTR), the start AUG (sAUG) at the beginning of the protein encoding open reading frame (ORF) terminated by a stop codon (UAA, UAG, or UGA), a 3′UTR, and a poly(A) tail. The 5′ cap is bound by the cap-binding complex eIF (eukaryotic initiation factor) 4F, consisting of eIF4E, 4A, and 4G. Secondary structures can be present in the 5′UTR by C-G, A-U, and G-U base pairing. These structures are unwound by the RNA helicase eIF4A. Through this action a single-stranded landing platform is created for the 43S pre-initiation complex, comprising the 40S ribosomal subunit, eIF2-GTP, and Met-tRNA. This complex scans along the

5′UTR, normally 60–100 nucleotides long, to the sAUG (17). The 60S ribosomal subunit joins, and protein synthesis begins. This mechanism for translation initiation is called cap-dependent ribosomal scanning and satisfactorily explains initiation on the vast majority of mRNAs. In this process, several determinants influence the efficiency of translation initiation, such as secondary structure and AUG codons in the 5′UTR, the nucleotide sequence surrounding the sAUG (18), and also the 3′UTR and poly(A) tail. A few cellular mRNAs and some viral RNAs use a cap-independent mechanism for initiation. These often very complex 5′ UTRs harbor an internal ribosomal entry site (IRES) in their 3′ part, able to attract ribosomal subunits independent of the 5′ end (26,41).

The 3′UTR and poly(A) tail greatly influence mRNA stability. However, during embryonic development both the 3′UTR and mainly the length of the poly(A) tail also determine translational efficiency by controlling mRNA localization and translational activation or repression (10,27,45,46).

All known aspects influencing the translatability and stability of the message will be the focus of this paper, by presenting experimental and literature data. The experiments deal with transfection of Cos-1 cells and injection of embryos of the clawed frog *Xenopus laevis*. Furthermore, recommendations regarding vector design are provided to assist in constructing a vector, taking care of special topics such as protein synthesis and mRNA stability.

MATERIALS AND METHODS

Construction of the Plasmids

The T_7TS plasmid for in vitro transcription (24) contains the *Xenopus* β-globin 5′UTR, sites for the insertion of an ORF, followed by the globin 3′UTR and a track of 30 A and C residues. The GFP ORF was inserted as an *Nco*I-*Bam*HI GFP fragment (pEGFP®; Clontech Laboratories, Palo Alto, CA, USA). The $A_{30}C_{30}$ tail was deleted by creating a *Pst*I site at the end of the 3′UTR by PCR. Digestion with *Pst*I and re-ligation removes the $A_{30}C_{30}$ tail, resulting in T_7TS-glob-GFP(-tail) (gift from G.C. Scheper, Dundee, UK). The T_7TS-HRV/EMCV/FMDV-CAT plasmids were constructed by inserting the 5′UTR-CAT sequences as *Eco*RV-*Bam*HI fragments into a blunt-ended *Hin*dIII/*Bgl*II-digested T_7TS plasmid. The HRV fragment contained the 3′ 592 bp of the 5′UTR (32), the EMCV 5′UTR was a 3′ 607-bp fragment (22), and the FMDV 5′UTR was a 530-bp fragment described earlier (39). The PV-CAT sequence (26) was inserted as a *Hin*dIII-*Bam*HI fragment into the *Hin*dIII/*Bgl*II-digested T_7TS, resulting in T_7TS-PV-CAT.

CS2-GFP was generated by inserting the GFP ORF as a *Bam*HI-*Xba*I fragment (pEGFP) into the *Bam*HI/*Xba*I-digested pCS2+ (40). CS2-glob-GFP was constructed by inserting the globin 5′UTR-GFP sequence (*Hin*dIII-*Eco*RV fragment, from T_7TS-glob-GFP) into a modified CS2+, in which the *Hin*dIII site was cloned directly at the start site for transcription.

In Vitro Transcription

T_7TS-glob-GFP was linearized with *Eco*RV (behind the ORF) or *Hin*cII (behind the AC tail), T_7TS-glob-GFP(-tail) with *Hin*cII, and T_7TS-PV/HRV/EMCV/FMDV-CAT with *Sma*I. Capped transcripts were made using the T_7-MEGAscript™ kit (Ambion, Austin, TX, USA) in the presence of 7.5 mM m^7GpppG, ATP, CTP, UTP, and 1.5 mM GTP. After DNase treatment, transcripts were separated from the free nucleotides and cap-analogue by gel-filtration of the phenol-treated transcription mixture using a Pasteur pipet filled with Sephadex® (Amersham Pharmacia Biotech) G50-fine in TNE (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA), followed by ethanol precipitation.

In Vitro Translation

A *Xenopus* in vitro translation system was prepared by washing stage 8 embryos twice with excess of 10 mM HEPES, 100 mM KAc, 1 mM $MgAc_2$, 0.05 mM EDTA, and 5 mM DTT. Subsequently, they were suspended in the same buffer (5 µL/embryo) supplemented with 10 µg leupeptin and 1 µg pepstatin/mL. The extract was centrifuged three times for 10 min at 15 000× *g*, 4°C to remove insoluble material. This extract was aliquoted, snap-frozen in liquid N_2, and kept at -80°C. For in vitro translation for 90 min at room temperature, 6 µL of this extract were supplemented with 4 µL containing: 0.125 mM amino acids, 0.8 µg tRNA, 3 µCi [^{35}S]methionine (1300 Ci/mmol), RNase inhibitor, an energy reconstitution system (30), 150 mM KAc, 6 mM $MgAc_2$, and 80 ng RNA. Translation in homemade rabbit reticulocyte lysate was described elsewhere (30). Translation products were analyzed by SDS-PAGE and autoradiography.

DNA Transfection, Extract Preparation, CAT Assay, and GFP Analysis

Cos-1 cells were transfected with the GFP DNA constructs as a calcium-phosphate precipitate. After 24 h, cells were harvested from the culture dish using 10 mM EDTA in PBS and lysed by repeated freeze-thawing in 20 mM Tris-HCl (pH 7.6), 100 mM KCl, 10% glycerol, 0.1 mM EDTA, 7 mM β-mercaptoethanol (embryos were lysed similarly). After centrifugation at 15 000× *g* rpm for 5 min at 4°C, supernatants were transferred to a new tube. CAT activity was assayed as described previously (28). The amount of GFP was measured in PBS using a spectrofluorometer [Photon Technology International, Ford, West Sussex, UK; λ(excitation) = 470 nm, λ(emission) = 508 nm].

Embryo Manipulation and Micro-Injection

In vitro fertilized *Xenopus* eggs were dejellied using 2% cysteine (pH 7.8) and kept in 25% MMR (modified amphibian Ringers' solution: 25 mM NaCl, 0.5 mM KCl, 0.25 mM $MgSO_4$, 0.5 mM $CaCl_2$, 1.25 mM HEPES-NaOH, pH 7.8, 0.025 mM EDTA). Embryos were injected in 25% MMR containing 3% Ficoll® with 1 ng RNA. After 4–6 h, the medium was changed to 25% MMR, and the embryos were allowed to develop at 18°C.

Northern Blotting

RNA (10 µg, glyoxylated in the presence of DMSO) was separated by

electrophoresis in a 1.5% agarose gel containing 15 mM of phosphate (7.5 mM Na_2HPO_4 and 7.5 mM NaH_2PO_4) and transferred to a nylon filter (Hybond®-N). After UV cross-linking, filters were hybridized overnight with a randomly primed ^{32}P-labeled CAT or GFP probe in 180 mM Na_2HPO_4, 70 mM NaH_2PO_4, 1 mM EDTA, 1% BSA (fraction V), and 7% SDS at 65°C. After autoradiography and quantification, filters were stripped in boiling 0.1% SDS and subsequently hybridized with a *Xenopus* Histon-3 probe as a control for RNA loading.

RESULTS AND DISCUSSION

Cap-Structure

Translation initiation starts at the 5′ end of the message, by building a preinitiation complex at the cap. The vast majority of cellular mRNAs absolutely depend on the presence of a cap for their translation in vivo.

After transfection or injection of DNA, capping is not a concern, as the addition of a 5′ terminal guanine and its methylation automatically occurs subsequently after transcription initiation by guanylyl transferase and methyltransferases. Experimental data indicate that RNA injection requires capped mRNA (47), which can be synthesized in vitro by adding an excess of m^7GpppG over GTP to the transcription assay.

Picornavirus 5′UTRs are naturally uncapped and initiate cap-independently using an IRES. We have investigated whether the cap-structure is also required in *Xenopus* injection studies when the mRNA contains an IRES. It is known that the Mengo IRES, closely related to encephalomyocarditis virus (EMCV), is recognized in *Xenopus* embryos, whereas the poliovirus (PV) IRES is not (8). We generated capped and uncapped RNAs containing the PV or EMCV IRES preceding the CAT ORF. These transcripts were injected into the one-cell stage *Xenopus* embryo, and embryo extracts were assayed for CAT activity at developmental stage 8 (Figure 1A). Both uncapped transcripts were very inefficiently translated, whereas only the capped EMCV-CAT mRNA yielded high CAT activity. The northern blot in Figure 1B shows that the uncapped transcripts were very unstable in the embryo, explaining their low translation into protein. This indicates that, in spite of the ability of the uncapped EMCV IRES RNA to recruit ribosomes, this process does not prevent mRNA degradation. From these data we conclude that injected transcripts, even with an IRES, need to be capped for high stability and expression.

Capped RNA for injection is made by transcription in the presence of millimolar amounts of cap-analogue (see Materials and Methods). Injection of micromolar amounts of cap-analogue results in embryonic death, as the analogue efficiently binds eIF4E and thereby severely inhibits protein synthesis. We routinely use gel filtration with Sephadex G50-fine of deproteinized RNA, which efficiently eliminates the cap-analogue in contrast to phenol extraction followed by ethanol precipitation.

We have developed an in vitro system derived from *Xenopus* embryos to test uncapped mRNAs. This is not helpful in protein overexpression studies but is useful to detect IRES activity. A crude concentrated translation extract can be made within one hour from *Xenopus* embryos of the desired stage. Figure 2 shows translation of uncapped PV, human rhinovirus (HRV), EMCV, and foot-and-mouth disease virus (FMDV)-CAT mRNAs in such a *Xenopus* lysate and in rabbit reticulocyte lysate. As is already known from literature (21), the PV IRES is recognized with a lower efficiency than the EMCV and FMDV IRESs in non-supplemented reticulocyte lysate (Figure 2A). Neither the PV IRES nor the HRV IRES was recognized by the in vitro *Xenopus* system, whereas the EMCV and FMDV IRESs were efficiently recognized in this in vitro assay (Figure 2B). The PV and EMCV results correspond with the injection experiment (Figure 1), showing the reliability of this *Xeno-*

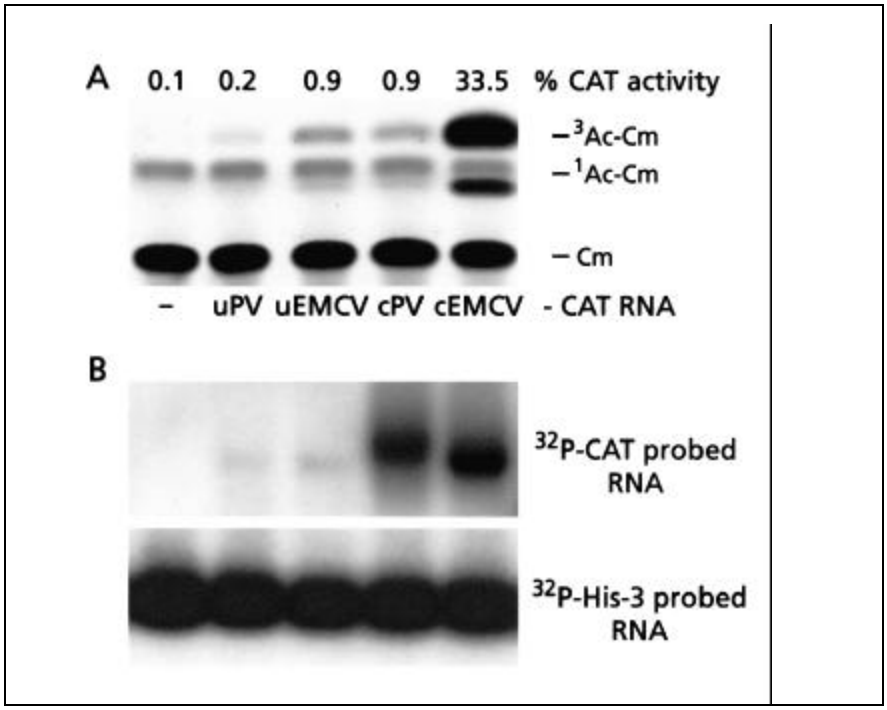

Figure 1. Translation and stability of capped and uncapped IRES-containing transcripts in *Xenopus* embryos. Fertilized *Xenopus* eggs were injected with in vitro synthesized capped (c) or uncapped (u) CAT mRNAs containing the PV or EMCV IRES. Embryo extracts were prepared at developmental stage 8, and total RNA was isolated. Extracts were tested for CAT activity; extract from uninjected embryos was used as a negative control (-). The percentage of chloramphenicol (Cm) converted to the acetylated forms ($^{1/3}$Ac-Cm) is indicated as % CAT activity (A). The band between the two acetylated forms is a degradation product, also seen in the absence of CAT-activity. Total RNA was assayed by northern blotting, first using a randomly primed ^{32}P-labeled CAT probe, followed by a *Xenopus* Histon-3 probe (B).

pus extract. As known, FMDV RNA contains two active initiation codons, leading to two CAT forms (Figure 2); the slower migrating form co-migrates with an endogenous protein synthesized in the *Xenopus* extract. Also HRV-RNA translation can initiate at an upstream AUG in vitro (2), explaining the two CAT forms (Figure 2A). With this assay, a potential *Xenopus* IRES can be tested for activity in a homologous system using an uncapped mRNA. As of now, no IRES-containing *Xenopus* 5′UTRs are known, but it is expected that they will be discovered in the near future.

5′UTR

Scanning ribosomal subunits are hampered by stable structures and AUG codons present in the 5′UTR. A number of reports indeed show that increasing the secondary structure and uAUG content in a 5′UTR severely affects translational efficiency (14,16,18,25). Of course, the presence of a uAUG may cause the synthesis of an elongated isoform of the protein, when the uAUG is in frame with the main ORF; because of leaky scanning, the ribosome can also reach the authentic start AUG, resulting in a second protein.

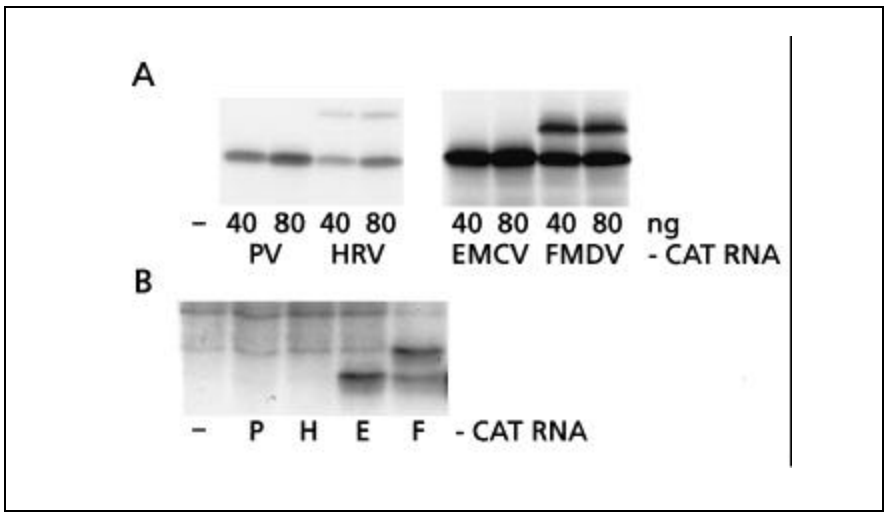

Figure 2. Efficiency of IRES-containing uncapped mRNAs in a *Xenopus*-derived translation lysate and in rabbit reticulocyte lysate. In vitro synthesized uncapped PV (P), HRV (H), EMCV (E), and FMDV (F)-CAT mRNAs were translated at a concentration of 40 and 80 ng/10 µL in reticulocyte lysate (A) and at 80 ng/10 µL in *Xenopus* extract (B). Details on extract preparation and translation conditions are described in Materials and Methods. ^{35}S-labeled CAT protein was visualized by autoradiography.

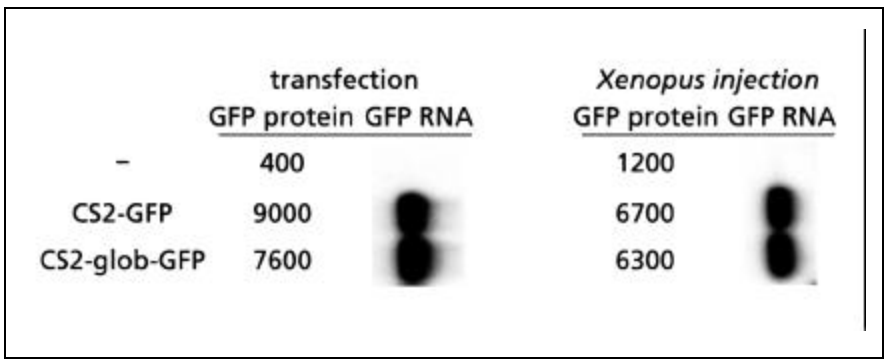

Figure 3. Comparison of the efficiency of the natural β-globin and a synthetic 5′UTR in tissue culture and *Xenopus* embryos. The CS2-GFP and CS2-glob-GFP DNA constructs were either transfected to Cos-1 cells or injected into one-cell stage *Xenopus* embryos. Protein extracts and total RNA were isolated 30 h after transfection or at developmental stage 18 in *Xenopus* development. The amount of GFP in the extracts was determined using a spectrofluorometer, and RNA was analyzed by northern blotting using a randomly primed ^{32}P-labeled GFP probe.

The average human 5′UTR has a length of 120 nucleotides, is devoid of secondary structure, and about 25% of all cDNAs analyzed contain a uAUG (37). This analysis contrasts with an analysis from 1987 of 700 vertebrate mRNAs, of which 9% contained a uAUG (15); this figure may have been skewed in favor of the simpler mRNAs. A number of regulatory proteins, such as growth factors (FGF-2, PDGF2, IGFII, EGF, and TGFβ), proto-oncogene products (c-mos, c-myc, c-syn, and c-bcl-2), receptors (RARβ2), and poly(A)-binding protein, are encoded by mRNAs with long, structured, AUG-containing 5′UTRs. A number of these 5′UTRs strongly suppress translation; furthermore, they can direct specific translational regulation (9,41). Most studied proteins are regulatory, and as the RT-PCR techniques rapidly improve, more and more cDNAs are found that contain complex 5′UTRs. This complexity can reside in the length of the 5′UTR, the CG content, the potential to fold into stable structures, and the presence of multiple uAUG codons.

In creating a vector for overexpression studies, two options can be used to eliminate the inhibitory or regulatory elements from a 5′UTR. (*i*) The major part of the 5′UTR can be removed by choosing a restriction enzyme that leaves an unstructured, non-AUG-containing part of the 5′UTR. However, to avoid chimeric 5′UTRs with unknown properties, the cleanest option is (*ii*) to amplify the ORF and to add suitable restriction sites at both ends by high-fidelity PCR. The resulting fragment can then be cloned into a commercially available or laboratory-optimized expression vector behind a natural or synthetic 5′UTR with proven efficiency.

An often-used expression vector, both for transfection and *Xenopus* injection is CS2+, designed by Turner and Weintraub (40). It contains a rather undefined sequence of 78 bp, without AUG codons and palindromic sequences, between the CMV promoter and polylinker. Cloning of a 5′UTR into the polylinker would result in a chimeric 5′UTR. We have tested whether this non-natural sequence acts itself as an efficient 5′UTR. For this purpose, we cloned the GFP ORF behind these 78 bp. For comparison the efficient β-globin 5′UTR was linked to GFP, precisely at the transcription start site. These two constructs were transfected and injected into *Xenopus* embryos, and the amounts of GFP protein and RNA were determined (Figure 3). Given the almost equal amounts of GFP RNA and protein, it can be concluded that this synthetic 5′UTR can serve as an efficient 5′UTR in embryos and in tissue culture. Insertion of a 5′UTR in CS2+ is unnecessary and even unadvisable.

Based on literature and experimental data, we suggest the use of a naturally occurring short 5′UTR. When a specific vector is required containing a synthetic 5′UTR, it needs to be carefully checked for palindromic sequences and AUG codons (for example, present in the *Nco*I, *Sph*I, *Nsi*I, *Bbu*I, *Nde*I, and *Sty*I restriction sites).

Sequence Surrounding the sAUG (Kozak Context)

Arrest of the scanning ribosomal subunit at the sAUG is mediated by codon-anticodon recognition between the initiator Met-tRNA and the sAUG. The sequence surrounding the sAUG is important in this process, with a consensus sequence for higher vertebrates of GCCA/GCCAUGG (15). The most highly conserved nucleotides within this consensus are the purine, usually an A at -3 and the G at +4; if these nucleotides are present, then the others only marginally contribute. Mutations affecting A^{-3} or G^{+4} strongly impair sAUG recognition in vivo and in vitro, resulting in ribosomal scanning beyond the sAUG and eventually initiation at a downstream AUG (18,19). Different species have a different consensus sequence. Therefore, applying the abovementioned consensus for expression in lower or invertebrate cells, as is often seen, may lead to lower expression levels. The consensus sequences for various species can be found in the TransTerm database (6); it gives, for example, A/GCCATGG for *Homo sapiens*, AA/CA/CATGG for *Xenopus laevis*, and AAA/CATGN for *Drosophila melanogaster*.

The ACC sequence upstream of the sAUG can always be generated by PCR. The G^{+4} can pose a problem, as this could alter the first amino acid of the protein, and thereby its stability (43) and activity. Fortunately, analysis of 699 vertebrate mRNAs revealed that as many as 46% contain the G at position +4 (15). Ideally, an expression vector suitable for human and *Xenopus* studies contains an efficient 5′UTR, an A residue at -3, directly followed by a unique *Nco*I site (CCATGG). The ORF with an *Nco*I site overlapping the sAUG can then be cloned in the *Nco*I site.

Protein-Encoding ORF

Using RNA, the sequence of the ORF from the start to the stop codon should obviously be without intron sequences. Although some expression vectors contain stop codons in the three reading frames after the polylinker, it is advisable to clone the ORF cDNA with its own stop codon, and preferably with the nucleotide behind the stop codon. Similarly as for the initiation context, the termination region also has a species-specific consensus, which can be found in the TransTerm database (6). Especially, the nucleotide behind the stop codon (+4 position) is well conserved and is involved in the efficiency of translation termination (35). In human and *Drosophila* genes, the G and A are most often present, whereas an A is preferred in *Xenopus* and invertebrates.

Cloning an ORF with its own termination region also assures that the resulting protein has the original C-terminus. This can be important for the activity of the protein, for its targeting to the peroxisome (1), plasma membrane (7), or trans-Golgi network (13).

3′UTR and Poly(A) Tail

The 3′UTR can influence the translatability of the mRNA by controlling RNA masking (34), influencing polyadenylation [e.g., cyclins and c-mos (48)], protein binding [e.g., caudal and hunchback (28,33)], and mRNA localization. Especially during embryonic development, there is a wealth of examples of 3′UTRs and polyadenylation events influencing translation efficiency (9,31,46). This type of regulation is not restricted to embryogenesis and is also seen for lipoprotein lipase RPE65

and 15-lipoxygenase expression (12, 20,38). Furthermore, adenylate/ uridylate-rich elements, which are found in the 3′UTR of many proto-oncogenes, transcription factors, and cytokines, influence mRNA stability (4).

These data indicate that cloning a gene with its own 3′UTR, even in truncated form, can interfere with efficient, unregulated protein expression. We have not investigated the effect of inserting various sequences between the ORF and polyadenylation signal. In our experience, however, part of the CAT or luciferase 3′UTR, the β-globin 3′UTR, and the vector sequence present in CS2+ as 3′UTR do not interfere with expression in transfection or injection experiments. Therefore, we recommend the above-mentioned sequences as good candidates for a non-regulated efficient 3′UTR.

Expression vectors often contain a strong viral polyadenylation signal; the late simian virus 40 (SV40) terminator is most frequently used and is also present in CS2+. This signal results in one species of polyadenylated stable RNA (Figure 3). In several baculovirus expression vectors, the SV40 early terminator is present. Recently, it was shown in insect cells that compared to the original p10 3′UTR the early terminator results in lower mRNA levels and reduced protein expression (42). This suggests that in this system a heterologous polyadenylation signal is less efficient than the homologous one.

The most frequently used vector for the generation of transcripts for *Xenopus* injection is pT_7TS, developed by Patterson and Krieg (24), containing the *Xenopus* β-globin 5′ and 3′UTRs, and a track of 30 A and C residues. RNA transcribed from this vector is much more efficient in the embryo than, for example, Bluescript® RNA without tail (44). We have determined the effects of removing the AC tail (glob-GFP-3′UTR) and the 3′UTR (glob-GFP), and replacing the AC tail for 75 A residues. Three capped transcripts—glob-GFP, glob-GFP-3′UTR, and glob-GFP-3′UTR-AC—were injected into *Xenopus* embryos, and their stability and translatability were determined at two developmental stages (Figure 4). At stage 5, the three mRNAs displayed similar stabilities, whereas the translational efficiency strongly depended on

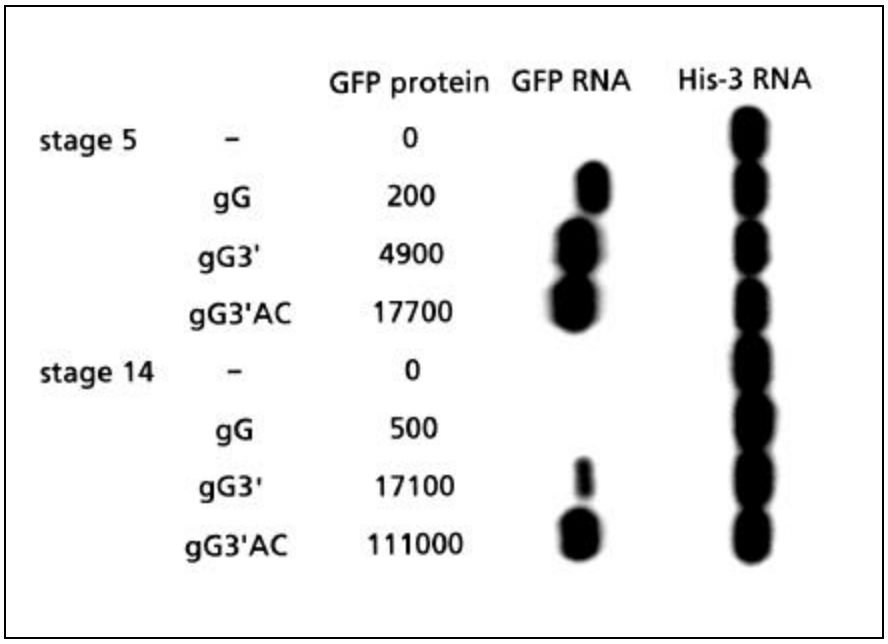

Figure 4. Influence of 3′ sequences on the stability and translatability of injected mRNAs. In vitro synthesized capped transcripts were injected into fertilized *Xenopus* eggs: glob-GFP (gG), glob-GFP-3′UTR (gG3′), and glob-GFP-3′UTR-AC (gG3′AC). Embryo extracts were prepared at developmental stages 5 and 14, and the amounts of accumulated GFP were determined. Data were corrected for the background signal of uninjected embryos (-). Total RNA was isolated at the same stages and assayed by northern blotting using a ^{32}P-labeled GFP probe, followed by a Histon-3 probe.

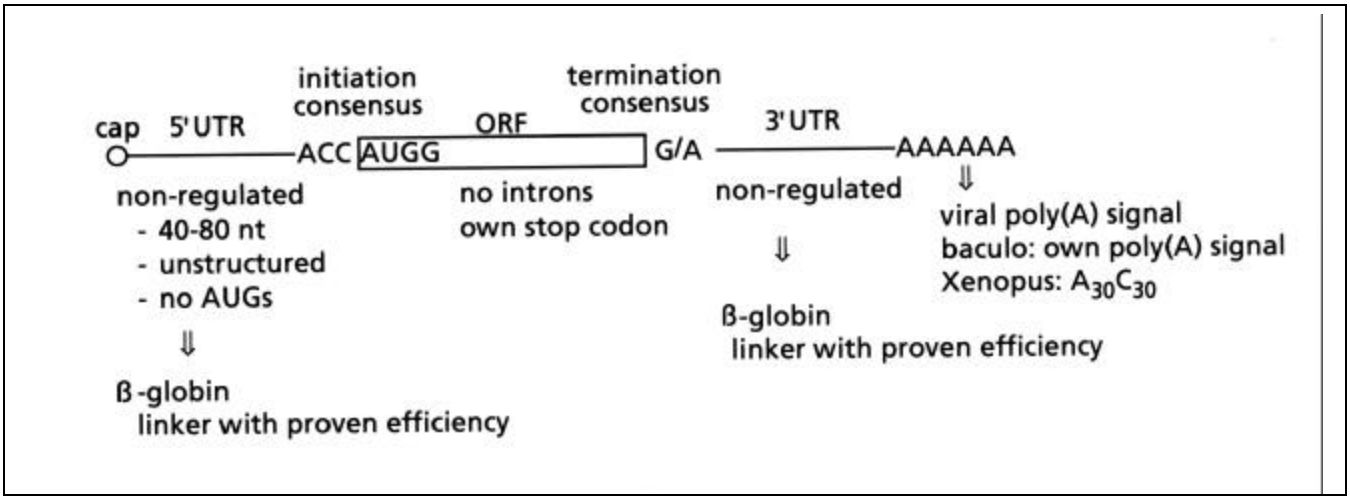

Figure 5. Schematic representation of a stable mRNA with high translational potential. See text for details.

the 3′ sequences; the RNA without 3′UTR and AC tail was translated with a 90-fold lower efficiency than the longest RNA. This effect was even more pronounced at stage 14, when both mRNA stability and translatability interfere with expression. The northern blot shows that the longest RNA was hardly degraded from stage 5 to 14, resulting in GFP accumulation linearly in time. The 3′ truncated RNAs were more labile; hardly any glob-GFP RNA was present anymore at stage 14, explaining the very modest increase in GFP. From these results, it can be concluded that both the 3′UTR and the AC tail are essential for stability, as well as for translatability in *Xenopus* embryos. Changing the AC tail for an A tail of 75 residues did not have an effect on the RNA stability, nor on translation efficiency (results not shown); therefore, we do not see an advantage in using polyadenylated mRNA in injection experiments.

At first sight, it is difficult to imagine how a 3′UTR and A tail influence translation initiation at the 5′ end of the message. Recently, it was suggested that the 3′ and 5′ ends of the mRNA are in close proximity because of circularization. In yeast and in mammals, this mRNA circularization is accomplished by poly(A)-binding protein, which can interact with both the poly(A) tail and eIF4G, part of the cap-binding eIF4F complex (11,29). Bridging the mRNA termini might promote re-initiation of terminating ribosomes or might efficiently attract new ribosomal subunits, both resulting in enhanced translation.

CONCLUSION

We have discussed the presently known aspects in a mature mRNA important for efficient translation and mRNA stability. Literature and experimental data were predominantly based on transfection of Cos-1 cells and injection of *Xenopus* embryos. Several literature examples can be found in which cell-specific translation was observed with picornaviral IRESs (3) and cellular IRESs, as described for c-myc (36) and FGF-2 (5). This cell specificity must reside in the untranslated regions, as reporter genes were used in these studies. Nonetheless, we assume that the general ideas about constructing an expression vector are also applicable for other cell lines and embryos. In Figure 5, we summarize the mentioned suggestions by presenting the construction of a stable mRNA with high translational potential.

ACKNOWLEDGMENTS

We thank Marcelle Kasperaitis and Kirsten van Nierop for excellent technical assistance, Olivier Destrée (Hubrecht Laboratory for Developmental Biology, The Netherlands) for support with the embryo injection experiments, and Gert Scheper (Dundee, UK) for providing the T_7TS-glob-GFP(-tail) construct.

REFERENCES

1. **Amery, L., C. Brees, M. Baes, C. Setoyama, R. Miura, G.P.Mannaerts, and P.P. Van Veldhoven.** 1998. C-terminal tripeptide Ser-Asn-Leu (SNL) of human D-aspartate oxidase is a functional peroxisome-targeting signal. Biochem. J. *336*:367-371.
2. **Borman, A. and R.J. Jackson.** 1992. Initiation of translation of human rhinovirus RNA: mapping the internal ribosome entry site. Virology *188*:685-696.
3. **Borman, A.M., P. Le Mercier, M. Girard, and K.M. Kean.** 1997. Comparison of picornaviral IRES-driven internal initiation of translation in cultured cells of different origins. Nucleic Acids Res. *25*:925-932.
4. **Chen, C.Y. and A.B. Shyu.** 1995. AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem. Sci. *20*:465-470.
5. **Créancier, L., D. Morello, P. Mercier, and A.-C. Prats.** 2000. Fibroblast growth factor 2 internal ribosome entry site (IRES) activity ex vivo and in transgenic mice reveals a stringent tissue-specific regulation. J. Cell Biol. *150*:275-281.
6. **Dalphin, M.E., C.M. Brown, P.A. Stockwell, and W.P. Tate.** 1998. The translational signal database, TransTerm, is now a relational database. Nucleic Acids Res. *26*:335-337.
7. **Fernandez-Larrea, J., A. Merlos-Suarez, J.M. Urena, J. Baselga, and J. Arribas.** 1999. A role for a PDZ protein in the early secretory pathway for the targeting of proTGF-alpha to the cell surface. Mol. Cell *3*:423-433.
8. **Gamarnik, A.V. and R. Andino.** 1996. Replication of poliovirus in *Xenopus* oocytes requires two human factors. EMBO J. *15*:5988-5998.

9. **Gray, N.K. and M. Wickens.** 1998. Control of translation initiation in animals. Ann. Rev. Cell Dev. Biol. *14*:399-458.
10. **Hake, L.E. and J.D. Richter.** 1997. Translational regulation of maternal mRNA. Biochim. Biophys. Acta *1332*:M31-M38.
11. **Imataka, H., A. Gradi, and N. Sonenberg.** 1998. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. EMBO J. *17*:7480-7489.
12. **Kern, P.A., G. Ranganathan, A. Yukht, J.M. Ong, and R.C. Davis.** 1996. Translational regulation of lipoprotein lipase by thyroid hormone is via a cytoplasmic repressor that interacts with the 3′ untranslated region. J. Lipid Res. *37*:2332-2340.
13. **Kjer-Nielsen, L., C. Van Vliet, R. Erlich, B.H. Toh, and P.A. Gleeson.** 1999. The Golgi-targeting sequence of the peripheral membrane protein p230. J. Cell Sci. *112*:1645-1654.
14. **Kozak, M.** 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA *83*:2850-2854.
15. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. *15*:8125-8148.
16. **Kozak, M.** 1989. Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic mRNAs. Mol. Cell. Biol. *9*:5134-5142.
17. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. *108*:229-241.
18. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. *266*:19867-19870.
19. **Kozak, M.** 1997. Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. EMBO J. *16*:2482-2492.
20. **Liu, S.Y. and T.M. Redmond.** 1998. Role of the 3′-untranslated region of RPE65 mRNA in the translational regulation of the RPE65 gene: identification of a specific translation inhibitory element. Arch. Biochem. Biophys. *357*:37-44.
21. **Meerovitch, K., Y.V. Svitkin, H.S. Lee, F. Lejbkowicz, D.J. Kenan, E.K.L. Chan, V.I. Agol, J.D. Keene, and N. Sonenberg.** 1993. La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. J. Virol. *67*:3798-3807.
22. **Oudshoorn, P., A. Thomas, G. Scheper, and H.O. Voorma.** 1990. An initiation signal in the 5′ untranslated leader sequence of encephalomyocarditis virus RNA. Biochem. Biophys. Acta *1050*:124-128.
23. **Pain, V.M.** 1996. Initiation of protein synthesis in eukaryotic cells. Eur. J. Biochem. *236*:747-771.
24. **Patterson, K.D. and P.A. Krieg.** 1996. Overexpression of the tinman-related genes XNkx-2.5 and XNkx-2.3 in *Xenopus* embryos results in myocardial hyperplasia. Development *122*:3549-3556.
25. **Pelletier, J. and N. Sonenberg.** 1985. Insertion mutagenesis to increase secondary structure within the 5′ noncoding region of a eukaryotic mRNA reduces translational efficiency. Cell *40*:515-526.
26. **Pelletier, J. and N. Sonenberg**. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature *334*:320-325.
27. **Richter, J.D.** 1996. Dynamics of poly(A) addition and removal during development, p. 481-503. *In* J.W.B. Hershey, M.B. Mathews, and N. Sonenberg (Eds.), Translational Control. CSH Laboratory Press, Cold Spring Harbor, NY.
28. **Rivera-Pomar, R., D. Niessing, U. Schmidt-Ott, W.J. Gehring, and H. Jäckle.** 1996. RNA binding and translational suppression by bicoid. Nature *379*:746-749.
29. **Sachs, A.B., P. Sarnow, and M.W. Hentze.** 1997. Starting at the beginning, middle, and end: Translation initiation in eukaryotes. Cell *89*:831-838.
30. **Scheper, G.C., H.O. Voorma, and A.A.M. Thomas.** 1992. Eukaryotic initiation factors -4E and -4F stimulate 5′ cap-dependent as well as internal initiation of protein synthesis. J. Biol. Chem. *267*:7269-7274.
31. **Seydoux, G.** 1996. Mechanisms of translational control in early development. Curr. Opin. Genet. Dev. *6*:555-561.
32. **Skern, T., W. Sommergruber, D. Blaas, P. Gruendler, F. Fraundorfer, C. Pieler, I. Fogy, and E. Kuechler.** 1985. Human rhinovirus 2: complete nucleotide sequence and proteolytic processing signals in the capsid protein region. Nucleic Acids Res. *13*:2111-2126.
33. **Sonoda, J. and R.P. Wharton**. 1999. Recruitment of nanos to hunchback RNA by pumilio. Genes Dev. *13*:2704-2712.
34. **Spirin, A.S.** 1996. Masked and translatable messenger ribonucleoproteins in higher eukaryotes, p. 319-334. *In* J.W.B. Hershey, M.B. Mathews, and N. Sonenberg (Eds.), Translational Control. CSH Laboratory Press, Cold Spring Harbor, NY.
35. **Stansfield, I., K.M. Jones, and M.F. Tuite.** 1995. The end in sight: terminating translation in eukaryotes. TIBS *20*:489-491.
36. **Stoneley, M., T. Subkhankulova, J.P.C. Le Quesne, M.J. Coldwell, C.J. Jopling, G.J. Belsham, and A.E. Willis.** 2000. Analysis of the c-myc IRES; a potential role for cell-type specific trans-acting factors and the nuclear compartment. Nucleic Acids Res. *28*:687-694.
37. **Suzuki, Y., D. Ishihara, M. Sasaki, H. Nakagawa, H. Tata, T. Tsunoda, M. Watanabe, T. Komatsu et al.** 2000. Statistical analysis of the 5′ untranslated region of human mRNA using "Oligo-Capped" cDNA libraries. Genomics *64*:286-297.
38. **Thiele, B.J., M. Berger, A. Huth, I. Reimann, K. Schwarz, and H. Thiele.** 1999. Tissue-specific translational regulation of alternative rabbit 15-lipoxygenase mRNAs differing in their 3′-untranslated regions. Nucleic Acids Res. *27*:1828-1836.
39. **Thomas, A.A.M., R. Rijnbrand, and H.O. Voorma.** 1996. Recognition of the initiation codon for protein synthesis in foot-and-mouth disease virus RNA. J. Gen. Virol. *77*:265-272.
40. **Turner, D.L. and H. Weintraub.** 1994. Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. Genes Dev. *8*:1434-1447.
41. **van der Velden, A.W. and A.A.M. Thomas.** 1999. The role of the 5′ untranslated region of an mRNA in translation regulation during development. Int. J. Biochem. Cell Biol. *31*:87-106.
42. **Van Oers, M.M., J.M. Vlak, H.O. Voorma, and A.A.M. Thomas.** 1999. Role of the 3′ untranslated region of baculovirus p10 mRNA in high-level expression of foreign genes. J. Gen. Virol. *80*:2253-2262.
43. **Varshavsky, A.** 1996. The N-end rule: functions, mysteries, uses. Proc. Natl. Acad. Sci. USA *93*:12142-12149.
44. **Vize, P.D., A. Hemmati-Brivanlou, R.M. Harland, and D.A. Melton.** 1991. Assays for gene function in developing *Xenopus* embryos. Methods Cell Biol. *36*:367-387.
45. **Wickens, M., J. Kimble, and S. Strickland.** 1996. Translational control of developmental decisions, p. 411-450. *In* J.W.B. Hershey, M.B. Mathews, and N. Sonenberg (Eds.), Translational Control. CSH Laboratory Press, Cold Spring Harbor, NY.
46. **Wickens, M., P. Anderson, and R.J. Jackson.** 1997. Life and death in the cytoplasm: messages from the 3′ end. Curr. Opin. Genet. Dev. *7*:220-232.
47. **Wormington, M.** 1991. Preparation of synthetic mRNAs and analyses of translational efficiency in microinjected *Xenopus* oocytes. Methods Cell Biol. *36*:167-183.
48. **Wormington, M.** 1994. Unmasking the role of the 3′ UTR in the cytoplasmic polyadenylation and translation regulation of maternal mRNAs. Bioessays *16*:533-535.

Received 20 November 2000; accepted 2 April 2001.

Address correspondence to:
Dr. Adri A.M. Thomas
Department of Developmental Biology
Utrecht University
Padualaan 8
3584 CH Utrecht, The Netherlands
e-mail: a.a.m.thomas@bio.uu.nl