Ribosomes stalling on uORF1 in the Xenopus Cx41 5′ UTR inhibit downstream translation initiation

Hedda A. Meijer* and Adri A. M. Thomas

Department of Developmental Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received February 12, 2003; Revised and Accepted April 14, 2003

ABSTRACT

Translation of Xenopus laevis Connexin41 mRNA is strongly controlled by the three upstream open reading frames (uORFs) in its 5′ untranslated region. Mutation of uAUG1 into AAG induced a 100-fold increase in translation of a green fluorescent protein (GFP) reporter ORF. The termination codon of uORF1 was mutated and the uORF was linked in-frame with the GFP ORF, enabling visualisation of initiation at uAUG1 by synthesis of an elongated GFP form. Unexpectedly, hardly any elongated GFP was made, suggesting that translation of uORF1 in wild-type mRNA causes constraining of the entry of 40S ribosomal subunits upstream of uORF1. A rare leucine codon, the third codon of uORF1, contributed to the slow translation and thus to slow scanning. Replacement of the rare leucine codon in uORF1 with a common leucine codon stimulated GFP translation. Remarkably, the rare leucine codon, the termination codon of uORF1, uAUG2 and uAUG3 all improved recognition of uAUG1. Apparently, the block formed by a stalled ribosome on any element in uORF1 prevented the landing of new ribosomal subunits next to the cap and therefore downregulated GFP translation.

INTRODUCTION

The binding of the 40S ribosomal subunit to mRNA is mediated by eukaryotic initiation factors (eIFs). After landing of the 40S subunit next to the m7G cap structure, the mRNA is scanned from 5′ to 3′ for initiation codons, in general AUG (1). Recognition of the initiation codon is mediated by eIF1, eIF2 and base pairing of the anticodon of Met-tRNAi with the AUG codon (2), and is dependent on sequences surrounding the AUG codon. The most frequently found nucleotides surrounding the AUG start codon in Xenopus laevis mRNAs are MMAMMAUGR (M = A or C, R = A or G) (http://uther.otago.ac.nz/Transterm.html). The most frequent nucleotides (in bold) are A at −3 and A or G at +4 (the A of the AUG codon is designated as +1). These two nucleotides are considered the most important nucleotides next to the AUG codon involved in recognition. A good initiation site contains both these nucleotides. When only one of these nucleotides is present, the sequence is said to be adequate for initiation. Whether these most frequent nucleotides indeed correspond to the most active initiation sequence, as has been shown for higher eukaryotes (3), remains to be confirmed.

Many mRNAs have 5′ untranslated regions (UTRs) containing at least one AUG codon (uAUG) upstream of the initiation codon preceding the main open reading frame (ORF) (4–7). These uAUGs, especially those in a good context, will generally inhibit translation of the main ORF. However, several examples are described of uAUGs controlling the expression of the main ORF in a flexible way, dependent on the circumstances (6). When an uAUG is not recognised, the ribosome continues scanning, a process named leaky scanning (8). Constraining scanning could be a powerful mechanism to prevent harmful overproduction of potent regulatory proteins (7). A few examples exist in which uORFs stimulate expression of downstream ORFs (9–13).

Some mRNAs have complex 5′ UTRs, containing multiple uAUGs and sometimes overlapping uORFs, which seem to be crucial for correct expression. Xenopus laevis Connexin41 (Cx41) mRNA contains such a complex 5′ UTR. Expression of this mRNA is controlled by a nested set of three uORFs in the 174 nt 5′ UTR (14; Fig. 1). All three uAUGs are recognised by at least some ribosomes, since mutation of each of the individual uAUGs resulted in an increased translation of the reporter gene. The presence of uAUG1 had the largest effect on downstream initiation (14). The sequence around uAUG1 forms an adequate initiation site (U at −3 and G at +4, Fig. 1A). Although the 5′ UTR of Cx41 is at least partially translated (14), the designation 5′ UTR will be used throughout this study to indicate the sequence from cap site to initiation codon of the connexin mRNA.

The strong increase of translation after mutation of uAUG1 suggested that the majority of ribosomes that bound to the wild-type mRNA recognised this uAUG and dissociated from the mRNA after termination at uORF1. Translation of the main ORF was therefore expected to be dependent on the relatively few ribosomes that skipped the uAUGs. We tested this suggestion in this study by making constructs in which the termination codon of uORF1 was mutated. An extra nucleotide was inserted to link this uORF in-frame to the green fluorescent protein (GFP) ORF. These mutations allowed detection of initiation at uAUG1, in the form of an N-terminal extended GFP. The results show that a surprisingly low number of ribosomes initiate at uAUG1. Apparently, the
translation of uORF1 poses a severe block for scanning ribosomes. One of the contributing elements is a rare leucine codon, at the third position of uORF1. The results show a remarkably complex mechanism to control the frequency at which ribosomes reach the main ORF.

MATERIALS AND METHODS

Plasmid construction

All plasmids were derived from pT7TS and contain the (mutated) Cx41 5′ UTR, the GFP ORF, the X.laevis β-globin 3′ UTR, and a track of 30 A and 30 C residues. mRNAs with a poly(AC) tail have the same stability and are translated with the same efficiency as mRNAs with a poly(A) tail of 75 nt (15). The mutants used for translation analyses (Fig. 1C and Table 1) were based on the results presented in Figure 3. The #23 mutant was not injected and was only used for in vitro assays. N.D.: not done.

Figure 1. Translation of Cx41 is controlled by three uORFs. (A) Sequence of the X.laevis Cx41 5′ UTR. (GenBank™ accession no. AF238222). uAUG codons are underlined, termination codons of the uORFs are in italics, and the GFP AUG is in bold (nucleotides 171–173). (B) Schematic representation of the three overlapping uORFs in the Cx41 5′ UTR. uORF2 and uORF3 are in the same reading frame and share a termination codon, they are also in the same reading frame as the main ORF. uORF1 is in a different reading frame. (C) Schematic representations of the ORFs present in the different mutants used in this study. For a detailed description of these and other mutants, see Table 1. The relative translation efficiencies of the different mutants after injection into one-cell-stage Xenopus embryos are based on the results presented in Figure 3. The #23 mutant was not injected and was only used for in vitro assays. N.D.: not done.
温度为60°C。经过最后一个周期后，额外延长了10 min（72°C）。

所有PCR片段都进行了测序。对于两个构建体，D23和D23#1，GFP ORF被替换为氯霉素乙酰转移酶（CAT）ORF，使用NcoI限制位点在5’ UTR和ORF之间，以及在载体中的PstI位点（就在A30C30尾巴的下游）。

表1。用于翻译分析的Cx41 5’ UTR突变体

<table>
<thead>
<tr>
<th>Name</th>
<th>Template PCR</th>
<th>Primers PCR</th>
<th>Mutations</th>
<th>Position</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>--</td>
<td>5F1 5R1</td>
<td>CTAATGG→ CCATGG</td>
<td>174</td>
<td>Introduction NcoI site</td>
</tr>
<tr>
<td>Δ1</td>
<td>WT</td>
<td>5F2 5F1 5R1</td>
<td>uATG1→AAG</td>
<td>42</td>
<td>Annihilation of uAUG1</td>
</tr>
<tr>
<td>Δ23</td>
<td>WT</td>
<td>5F3 5F1 5R1</td>
<td>uATG2→AAG</td>
<td>98</td>
<td>Annihilation of uAUG2 and uAUG3</td>
</tr>
<tr>
<td>#1</td>
<td>WT</td>
<td>5F1 5R4</td>
<td>TAG1→TTG</td>
<td>126</td>
<td>uORF1 connected to ORF GFP</td>
</tr>
<tr>
<td>Δ23#1</td>
<td>Δ23</td>
<td>5F1 5R4</td>
<td>uATG2→AAG</td>
<td>98</td>
<td>uORF1 connected to ORF GFP</td>
</tr>
<tr>
<td>#23</td>
<td>WT</td>
<td>5F1 5R5</td>
<td>TAA→AAA</td>
<td>139</td>
<td>uORF2+3 connected to ORF GFP</td>
</tr>
<tr>
<td>CB±</td>
<td>WT</td>
<td>5F5 5R1</td>
<td>GCT→GCCG</td>
<td>46</td>
<td>Creation of rare codons for ALPV3 in uORF1</td>
</tr>
<tr>
<td>CB±</td>
<td>WT</td>
<td>5F6 5R1</td>
<td>TTA→CTG</td>
<td>47+49</td>
<td>Creation of common codons for ALPV3 in uORF1</td>
</tr>
<tr>
<td>#1CB±</td>
<td>#1</td>
<td>5F5 5R4</td>
<td>GCT→GCCG</td>
<td>46</td>
<td>Creation of rare codons for ALPV3 in uORF1</td>
</tr>
<tr>
<td>#1CB+</td>
<td>#1</td>
<td>5F6 5R4</td>
<td>TAG1→TTG</td>
<td>126</td>
<td>Creation of common codons for ALPV3 in uORF1</td>
</tr>
<tr>
<td>Δ23CB±</td>
<td>Δ23</td>
<td>5F5 5R1</td>
<td>GCT→GCCG</td>
<td>46</td>
<td>Creation of rare codons for ALPV3 in uORF1</td>
</tr>
<tr>
<td>Δ23CB+</td>
<td>Δ23</td>
<td>5F6 5R1</td>
<td>uATG2→AAG</td>
<td>98</td>
<td>Annihilation of uAUG2 and uAUG3</td>
</tr>
<tr>
<td>Δ23#1CB±</td>
<td>Δ23#1</td>
<td>5F5 5R4</td>
<td>GCT→GCCG</td>
<td>46</td>
<td>Creation of rare codons for ALPV3 in uORF1</td>
</tr>
<tr>
<td>Δ23#1CB+</td>
<td>Δ23#1</td>
<td>5F6 5R4</td>
<td>TAG1→TTG</td>
<td>126</td>
<td>Annihilation of uAUG2 and uAUG3</td>
</tr>
</tbody>
</table>

aPrimers are listed in Table 2.
bAll mutants are based on the WT Cx41 5’ UTR containing the NcoI restriction site.
cALPV = alanine, leucine, proline, valine, amino acid residues 2–5 of uORF1.
Table 2. Primers used for plasmid construction

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5'-3')abc</th>
<th>Position</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>5F1</td>
<td>CCGCAAGCTTGGCTCTCTATATCACACGC</td>
<td>1-18</td>
<td>HindIII</td>
</tr>
<tr>
<td>5F2</td>
<td>GGCACAATACGGAAATATATAGCTTACAGGAAG</td>
<td>23-58</td>
<td>Mutagenesis uATG1</td>
</tr>
<tr>
<td>5F3</td>
<td>CTACACAAATTTCTGGGCAWGGCAAWGCAGTGGGGGTTGCC</td>
<td>78-118</td>
<td>Mutagenesis uATG2+3</td>
</tr>
<tr>
<td>5F5</td>
<td>CCGCAAGCTTGGCTCTCTATATCACACACATATCGGAATTTATAGCTTACAGGAAG</td>
<td>1-78</td>
<td>HindIII, mutagenesis ALPVd into rare codons</td>
</tr>
<tr>
<td>5F6</td>
<td>CCGCAAGCTTGGCTCTCTATATCACACACATATCGGAATTTATAGCTTACAGGAAG</td>
<td>1-78</td>
<td>HindIII, mutagenesis ALPVd into common codons</td>
</tr>
<tr>
<td>5R1</td>
<td>CAGGCATACGGAAATATATAGCTTACAGGAAG</td>
<td>185-159</td>
<td>NcoI</td>
</tr>
<tr>
<td>5R4</td>
<td>TTAGCTCAACATATACGGAAATATATAGCTTACAGGAAG</td>
<td>178-108</td>
<td>NcoI, mutagenesis stop codon uORF1</td>
</tr>
<tr>
<td>5R5</td>
<td>TTAGCTCAACATATACGGAAATATATAGCTTACAGGAAG</td>
<td>178-101</td>
<td>NcoI, mutagenesis stop codon uORF2+3</td>
</tr>
</tbody>
</table>

abcRestriction sites are underlined.
W = A or T.
ALPV = alanine, leucine, proline, valine, amino acid residues 2-5 of uORF1.

Analysis of translation efficiency

Linearisation of DNA downstream of the A30C30 tail, in vitro transcription and in vitro translation assays were done as described (14). Theoretical capping efficiency is 80%; the un capped mRNAs do not contribute to protein synthesis as they are quickly degraded as well as poorly translated (15). Transcript yield and integrity were determined after Sephadex G50 gel filtration by spectrophotometry (1 A 260 U = 4 0 × 10^-15 g/ml or 26 nM). All constructs had the same molecular mass and were tested at equal molar concentration. Injection into one-cell stage X. laevis embryos and in vivo translation analyses by western and northern blot analysis were performed as described (14). All embryos were injected with 1 ng (2.6 fmol) of mRNA (in 10 nl). GFP antibody (Clontech) was used in a 100-fold dilution, CAT antibody (5 μl) in 60 nl of RIPA buffer. After addition of 8.5 μl of in vitro synthesised 35S-labelled protein, the mixture was incubated for 2 h at room temperature. Protein A–Sepharose with bound antibody and CAT protein was precipitated by centrifugation. The pellet was washed twice in high-salt buffer [500 mM NaCl, 1% (v/v) Triton X-100, 20 mM Tris–HCl pH 7.4] and twice in low-salt buffer [150 mM NaCl, 1% (v/v) Triton X-100, 20 mM Tris–HCl pH 7.4], and resuspended in Laemmli sample buffer. The unbound proteins in the supernatant were precipitated with 5% trichloroacetic acid. The resulting pellet was resuspended in Laemmli sample buffer complemented with 0.2 M Tris. Bound and unbound CAT protein was analysed by SDS–PAGE.

RESULTS

Efficiency of uAUG recognition

Mutation of the three uAUGs in the 5' UTR of Cx41 strongly increased translation of the GFP reporter (Fig. 1). The leaky scanning model suggests that the strong increase of GFP translation after removal of the uAUGs is due to efficient recognition of the uAUGs. To visualise initiation at uAUG1, the termination codon of uORF1 was mutated and the uORF was made in-frame with the GFP ORF (Table 1). Initiation at uAUG1 will result in the synthesis of an elongated GFP protein, designated #1 GFP.

Transcripts consisting of the (mutated) Cx41 5' UTR, the GFP ORF, the globin 3' UTR and an A30C30 tail were synthesised in vitro. In the Δ1 and Δ23 transcripts, uORF1 or uORF2 and uORF3 were mutated to AAG codons. The #1 transcript lacks the termination codon of uORF1, and the uORF was connected to the GFP ORF. To analyse the upstream initiation efficiency at uAUG1 in the absence of uORF2 and uORF3, Δ23#1 (no uORF2 and uORF3, uORF1 linked to the GFP ORF) was constructed.

Transcripts lacking initiation and/or termination codons in the 5' UTR were translated in a rabbit reticulocyte lysate for a qualitative analysis to determine whether the uAUGs were recognised in vitro and whether correctly sized proteins can be detected (Fig. 2). Translation of all transcripts resulted in the synthesis of correctly sized GFP. The nearly equal synthesis of GFP in all lanes was not unexpected, because earlier analysis in the in vitro reticulocyte system had shown that uAUG codons in the Cx41 5' UTR hardly influenced GFP translation in this assay (14). Mutation of the termination codon resulted nonetheless in the synthesis of elongated GFP, showing the recognition of uAUG1 in vitro.

Immunoprecipitation

Protein A–Sepharose CL-4B (Amersham Pharmacia Biotech) was suspended in RIPA buffer [20 mM Tris–HCl pH 7.5, 150 mM NaCl, 1% (v/v) Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 5 mM EDTA pH 8.0, 1 mM phenylmethylsulfonyl fluoride]. Per assay, 4 mg of protein A–Sepharose was incubated with the CAT antibody (5'–3'), 0.3 μl) in 60 μl of RIPA buffer for 1 h at room temperature. Unbound antibody was removed by washing with RIPA buffer, and the protein A–Sepharose was resuspended in 56.5 μl of RIPA buffer. After addition of 8.5 μl of in vitro synthesised 35S-labelled protein, the mixture was incubated for 2 h at room temperature. Protein A–Sepharose with bound antibody and CAT protein was precipitated by centrifugation. The pellet was washed twice in high-salt buffer [500 mM NaCl, 1% (v/v) Triton X-100, 20 mM Tris–HCl pH 7.4] and twice in low-salt buffer [150 mM NaCl, 1% (v/v) Triton X-100, 20 mM Tris–HCl pH 7.4].
Figure 2. In vitro analysis of the effect of mutation of the uORF1 termination codon. All transcripts contain the (mutated) Cx41 5′ UTR, the GFP ORF, the X.laevis β-globin 3′ UTR and a tail of 30 A and 30 C nucleotides. Transcripts (Fig. 1C and Table 1) were translated in a rabbit reticulocyte lysate in the presence of [35S]methionine, the resulting proteins were separated on a 15% 30:0.18 acrylamide/bisacrylamide gel, and visualised by autoradiography. Migration of GFP and elongated GFP (#1 GFP) is indicated on the right.

The same transcripts were analysed in vivo by injecting them into one cell-stage Xenopus embryos. After 24 h, pools of 10 embryos were made and proteins and RNA were isolated. Synthesised GFP protein was visualised by western blotting (Fig. 3A). The amount of eIF4A was used as a loading control (Fig. 3B). The amount of injected RNA was analysed by northern blotting, using a GFP probe, and found to be similar (Fig. 3C). After stripping, the blot was hybridised with a Xenopus histone 3 (XH3) probe to correct for the amount of RNA loaded (Fig. 3D). The translation efficiency of the different transcripts was calculated by dividing the amount of GFP protein by the amount of GFP mRNA, and both figures were corrected for the loading controls. The translation efficiency of the most efficiently translated transcript (#1) was set to 100; the other translation efficiencies were calculated relative to #1.

As also shown previously (14), mutation of the first uAUG (#1) strongly increased the translation of the GFP ORF (100-fold, Fig. 3E). This suggests that the majority of ribosomes that enter the mRNA translate uORF1 and are released after the termination codon. However, mutation of the termination codon of uORF1 did not result in large amounts of elongated GFP (indicated as #1 GFP). The amount of #1 GFP was on average 8% (see also Fig. 7F). Normal GFP translation was slightly enhanced with #1 mRNA compared with WT mRNA (3-fold after corrections, Fig. 3E), suggesting that prevention of termination induced 3-fold more 40S subunits to scan the entire 5′ UTR and reach the GFP ORF. Similar results were obtained with the Cx41 variant in which uAUG2 and uAUG3 were mutated to AAG. The inactivation of these uAUGs led to 7-fold more GFP, in agreement with earlier data (14). The removal of the termination codon in Δ23#1 led to higher GFP synthesis, as also observed with the #1 variant, suggesting that termination at uORF1 is a translational barrier.

The small amount of #1 GFP contrasts with the large amount of GFP in the Δ1 mutant. This might be due to a low affinity of the GFP antibody for the elongated #1 GFP form, which may therefore result in an underestimation of the initiation frequency at uAUG1. To exclude this effect, and to exclude the possibility that AUG recognition was dependent on the reporter sequence, the GFP ORF was replaced by the CAT ORF. The activity of two mutants, Δ23#1 and Δ23, was analysed by western and northern blotting, as described for the GFP constructs. The translation efficiency directed by the Δ23#1 and Δ23 transcripts was determined (Fig. 3F–J). The termination codon-mutated variant (Δ23#1) was twice as active as the Δ23 variant, just as has been found for the GFP constructs (Fig. 3E and J). Importantly, the upstream initiation efficiency measured as #1 GFP and #1 CAT was also similar (Fig. 3A and F).

Theoretically, the protein stability of the #1 GFP might be lower than that of GFP, resulting in an underestimation of the initiation frequency at uAUG1. This explanation is considered unlikely since the ratio of the elongated proteins to the normal forms of CAT and GFP is similar for both forms (Fig. 3A and F). Moreover, we will show later (see Fig. 7) that the amount of GFP and #1 GFP can be manipulated by changing the mRNA sequence, but keeping the amino acid sequence, and thus the protein stability, constant.

The amount of GFP and #1 GFP could not be directly assayed by immunoprecipitation with GFP antibodies (not shown). In contrast, the affinity of the CAT antibody for both forms could be measured in an immunoprecipitation assay. Analysis of the CAT products after in vitro translation of Δ23#1-CAT and Δ23-CAT transcripts in the presence of 35S-labelled methionine showed that the ratio of CAT isoforms in immunoprecipitated CAT was similar to the amount directly analysed. These controls indicate that the GFP and CAT antibodies were equally active towards the normal and the elongated forms, and thus that the initiation frequency at uAUG1 is quite low, in contrast to the expectation after analysis of the Δ1 transcript. In spite of the strong stimulation of GFP translation in the Δ1 variant, upstream initiation is apparently a rather rare event as measured by the amount of #1 GFP in the Δ23#1 variant. Removal of uAUG2 and uAUG3 induced even less ribosome initiation at uAUG1 in the #1 variant (Fig. 3A), suggesting that initiation at uAUG2 and uAUG3 influenced upstream initiation at uAUG1. This result resembles the increased initiation when a hairpin is positioned just downstream of an initiation codon (19).

Codon usage

The translation of uORF1 in the Cx41 5′ UTR, albeit rather rare, has apparently an enormous inhibitory effect on translation of the main ORF. This inhibitory effect was not caused by inefficient initiation at uAUG1 as only a minor portion of the ribosomes initiate at this uAUG. Therefore, the initiation, the elongation or the termination process at uORF1 might block scanning ribosomal subunits. The data in Figure 3 indicate that the termination codon of uORF1 indeed caused some delay for ribosomes heading towards the AUG of GFP, and thus that termination at uORF1 caused fewer ribosomes to reach the AUG of GFP. This effect (3-fold) is relatively minor compared with the 100-fold increase when uAUG1 is mutated. The total amount of initiations at the #1 transcript (visualised as GFP and #1 GFP together) is very small compared with the amount of #1 CAT in the Δ23#1 variant. Removal of uAUG2 and uAUG3 induced even less ribosome initiation at uAUG1 in the #1 variant (Fig. 3A), suggesting that initiation at uAUG2 and uAUG3 influenced upstream initiation at uAUG1. This result resembles the increased initiation when a hairpin is positioned just downstream of an initiation codon (19).

Downloaded from https://academic.oup.com/nar/article-abstract/31/12/3174/1395307 by guest on 07 January 2019
codons of uORF1 (Fig. 4A) represent the relative frequency with which the codon is used to encode the corresponding amino acid. When the codon for a particular amino acid is unique (as for tryptophan), this percentage is 100. Examples have been described in which codon bias can regulate translational efficiency (20). In eukaryotes, examples of this type of translational control are still rare. To test whether codon usage could indeed be responsible for the translational control by uORF1, we first used the rabbit reticulocyte assay. The reticulocyte lysate contains an amount of tRNA sufficient to drive translation to a level of 60%, compared with the normalised 100% level obtained after addition of sufficient exogenous tRNA (unpublished data).

We tested whether addition of tRNA to reticulocyte lysate would improve translation of the elongated #1 GFP more than GFP itself. As a control, a construct was used that synthesised three GFP forms, initiated at uAUG2, uAUG3 and the start AUG of GFP. The GFP ORF was replaced by the CAT ORF. Translation analysis was the same as described for the GFP transcripts except that for the western blot the GFP antibody was replaced by a CAT antibody. (K) The affinity of the CAT antibody for the two CAT forms is similar. After in vitro translation of the Δ23#1-CAT and Δ23-CAT transcripts in the presence of 35S-labelled methionine, the resulting proteins were precipitated by the same CAT antibody used for the translation analysis. Proteins were separated on an acrylamide/bisacrylamide gel prior to and after immunoprecipitation.

Figure 3. The uAUG1 codon is barely recognised in vivo. (A-E) One cell-stage Xenopus embryos were injected with GFP encoding transcripts containing the (mutated) Cx41 5′ UTR, as indicated above each lane. Protein and mRNA were isolated after 24 h. Protein derived from two embryos was separated on a 15% 30:0.18 acrylamide/bisacrylamide gel. The blot was divided into two parts; the lower part was incubated with an antibody against GFP (A), whereas the upper part was incubated with an eIF4A antibody as loading control (B). All lanes were from the same experiment and blot. A 5 μg aliquot of mRNA was glyoxylated and analysed on a sodium phosphate-buffered agarose gel. The blot was subsequently analysed by incubation with a GFP (C) and a Xenopus histone 3 (D) probe. GFP translation efficiency was calculated by dividing the amount of GFP protein by the amount of injected GFP mRNA (E). Both figures were corrected for the loading controls. (F–J) For two transcripts, the GFP ORF was replaced by the CAT ORF. Translation analysis was the same as described for the GFP transcripts except that for the western blot the GFP antibody was replaced by a CAT antibody. (K) The affinity of the CAT antibody for the two CAT forms is similar. After in vitro translation of the Δ23#1-CAT and Δ23-CAT transcripts in the presence of 35S-labelled methionine, the resulting proteins were precipitated by the same CAT antibody used for the translation analysis. Proteins were separated on an acrylamide/bisacrylamide gel prior to and after immunoprecipitation.
GFP. At the lower tRNA concentrations, initiation at uAUG1 was inhibited (Fig. 5C), whereas initiation at uAUG2 and uAUG3 was comparable with initiation at the main AUG (Fig. 5B). This result suggested indeed that sequences unique for uORF1 are responsible for the relatively inefficient translation, and that tRNA addition can reverse this inefficiency.

This encouraged us to construct a new set of mutants (Fig. 4B and C). In the first series, CB+, the rare leucine codon was mutated to the common CUG leucine codon that is three times more frequent in codon regions of *X.laevis* (http://www.kazusa.or.jp/codon). In the second series, CB−, the rare leucine codon was retained, and the surrounding codons for alanine, proline and valine were also mutated into rare codons. The codons downstream of the valine codon could not be mutated into rare codons, because all codons for these amino acids were used with similar frequency. The codon bias mutants were made in four different 5′ UTRs, the wild-type Cx41 5′ UTR, the D23 form, which is more efficient and allows better detection of changes in translation efficiency, the #1 form, allowing direct analysis of the codon bias on upstream initiation, and the D23#1 form, combining both purposes.

These CB− and CB+ mutants were transcribed and translated in vitro (Fig. 6). All variants produced similar amounts of GFP and #1 GFP. This result shows that the transcripts were of similar quality, allowing their use in vivo. The migration of #1 GFP synthesised from the D23#1 transcript is slightly changed compared with #1 GFP encoded by the #1 transcript due to the mutation of uAUG2 and uAUG3, which causes two amino acid changes in uORF1: C20S and C22S.

Transcripts were injected into *Xenopus* embryos, and protein and mRNA were isolated after 24 h to analyse the amount of produced GFP protein (Fig. 7A) relative to the amount of injected GFP mRNA present at the time of harvesting (Fig. 7C). Both figures were corrected for the loading controls (Fig. 7B and D). Since it is technically impossible to inject 12 transcripts in one experiment before the embryos start to divide, the transcripts were divided into four groups. Translation efficiency in each group was calculated relative to the transcript with the wild-type codon usage. The translation efficiencies are presented as the mean of five analyses (Figs 4C and 7E).

Replacement of the rare wild-type leucine codon with a common leucine codon in the wild-type Cx41 5′ UTR increased GFP translation 3.4-fold, indicating that this codon indeed delays progress of translation of uORF1. Addition of three rare codons next to the leucine codon reduced GFP translation below the wild-type level. The effect of codon

---

**Figure 4.** Analysis of codon usage in uORF1. (A) Sequence of the peptide encoded by uORF1. The numbers under the codons represent the frequency with which the codons are used to encode the corresponding amino acid in *X.laevis* ORFs (27). The rare leucine codon is in bold. (B) Codon usage of amino acids 2–5 in the codon bias (CB) mutants. The codons were made either minimal (CB−, four rare codons) or optimal (CB+, the rare leucine codon was replaced by a common leucine codon). Mutated nucleotides are underlined. (C) Translation efficiency of RNAs after injection into one cell-stage *Xenopus* embryos. For explanation of the different mutants, see Figure 1C. The relative translation efficiencies are calculated based on the results presented in Figures 3 and 7.

**Figure 5.** Effect of tRNA addition on uAUG recognition *in vitro*. Transcripts encoding different GFP forms, indicated above each panel, were translated in rabbit reticulocyte lysate complemented with 0–200 ng of tRNA per 5 μl assay. The resulting proteins were separated on a 15% 30:0.18 acrylamide/bisacrylamide gel. The gel was dried, and incorporated [35S]methionine was visualised by exposure to film (A). The bands were quantified. Percentage upstream initiation was calculated by dividing the amount of elongated GFP (#1 GFP or #23 GFP) by the total amount of GFP (B and C). The data were corrected for the number of methionines present in the different GFP forms. The open and closed circles represent two independent experiments.
usage was also analysed in the absence of uAUG2 and uAUG3. The stimulating effect of the removal of the rare leucine codon was less prominent, but the addition of the three rare codons repressed GFP translation strongly. A similar trend in GFP expression was also measured with the #1 constructs: removal of the rare UUA leucine codon increased GFP synthesis. Upstream initiation was visualised by the appearance of the #1 GFP protein after translation of the #1 transcripts. Exchange of the UUA for the CUG leucine codon effectively annihilated upstream initiation, an unexpected result. Apparently, the presence of the rare leucine codon improves translation initiation on the AUG codon just upstream, similar to the initiation-enhancing effect of a hairpin structure just downstream of an initiation codon (12).

**DISCUSSION**

uORF1 in the Cx41 5’ UTR strongly inhibited translation of the GFP reporter, as mutation of uAUG1 induced a 100-fold increase in GFP translation. This suggested a very efficient initiation at uAUG1, and consequently a very low number of ribosomes capable of initiation at the GFP AUG. The frequency of upstream initiation was measured by using a construct in which uORF1 was linked to the GFP ORF, resulting in an N-terminal elongated GFP form after initiation at uAUG1. In contrast to the large impact of uAUG1 on GFP translation, this codon was barely used for initiation. Therefore, translation of uORF1 must be a relatively slow event, preventing leaky scanning of the Cx41 5’ UTR by ribosomal subunits. The result implies that scanning ribosomes transverse the mRNA much faster than translating ribosomes.

Although initiation in general is the rate-limiting step in the translation of an ORF, some examples are described in which elongation or termination is the restrictive process. Translation of AdoMetDC is strongly dependent on the sequence of a peptide encoded by a small uORF. Elevated polyamines inhibit the termination of translation of the uORF by stabilising the complete nascent peptide, linked to the tRNA that decodes the final codon. This peptidyl-tRNA molecule interacts with a component of the translational apparatus. The resulting complex blocks upstream scanning 40S subunits and therefore inhibits translation (21). Termination at the second uORF of gpUL4 is a slow process because the peptidyl-tRNA complex is not released and this complex causes queuing of upstream scanning ribosomal subunits (22,23). Translation of Arg2 is dependent on a conserved uORF encoding the arginine attenuator peptide, AAP. The synthesis of this peptide also causes stalling, independently of the termination codon (24). Introduction of a rare leucine codon directly downstream of the removed termination codon resulted in an increased toe-print signal compared with the introduction of a common leucine codon. Remarkably, this extra stalling did not have any effect on downstream translation efficiency (24).

Mutation of the termination codon of uORF1 in the Cx41 5’ UTR caused a 3-fold increase in GFP translation, showing that the termination indeed delayed the arrival of ribosomes at the main ORF. Replacement of the rare leucine codon, the third codon of uORF1, by a common leucine codon also resulted in an increased GFP translation. The impact of mutation of uAUG1 is much stronger than the effect of the rare leucine codon and the termination codon. Therefore, additional delaying elements in uORF1 are expected. The leucine codon was the only rare codon in uORF1. Other candidates for a role in the slow translation of uORF1 are two tryptophan and two cysteine codons. Tryptophan is encoded by only one codon, whereas cysteine is encoded by two equally used codons. Both tryptophan and cysteine are infrequently used amino acids. The presence of two tryptophan codons and two identical cysteine codons in a uORF encoding a 28 amino acid peptide is remarkable.

Stalling of ribosomes can be shown by toe-printing: reverse transcription of polysomal RNA will terminate when 40S or 80S ribosomes block reverse transcription. Unfortunately, the reticulocyte data do not reflect the *Xenopus* data, and toe-printing proved impossible in *Xenopus* lysates: the lysates were only marginally active in translation, and the majority of ribosomes were not bound to mRNA.

The elimination of the rare leucine codon caused a less frequent initiation at uAUG1. Pausing of the elongating ribosome at the leucine codon apparently stimulated uAUG1 recognition. The block formed by the stalled ribosome might function comparably with a hairpin just downstream of an AUG codon (19). The optimal stimulatory effect of the hairpin was achieved when the distance between the AUG and the hairpin was 14 nt (19), corresponding to half the size of a 40S ribosomal subunit. The explanation of the stimulatory effect of a stalled ribosome at codon 3 of uORF1 is more complex. Only one codon separates the uAUG1 initiation codon from the rare leucine codon. The size of the elongating ribosome seems to exclude stalling of the scanning 40S subunit with the anticodon of Met-tRNA, placed opposite the uAUG codon. The distance between the 5’ cap structure and the rare leucine codon is 46 nt. A stalling ribosome might therefore prevent landing of a new 40S subunit next to the cap due to steric hindrance. Consequently, a new 40S subunit can land only when a rare leucine tRNA becomes bound to the ribosomal A-site and when the ribosome can continue elongation. A scanning 40S subunit, queued behind the elongating ribosome, will scan the 5’ UTR as fast as the elongating ribosome can incorporate the amino acids. This will result in slow scanning of the 5’ UTR. Apparently, this resulted in a more careful scanning and a better recognition of uAUG1. The rare leucine codon counterbalances the inefficient recognition of uAUG1 by delaying the translation of uORF1 and by improving the
Initiation frequency at uAUG1. Initiation at uAUG1 was stimulated by uAUG2 and uAUG3. Translation of uORF2 and uORF3 therefore might slow down upstream scanning ribosomal subunits in a way comparable with ribosomes translating uORF1. Probably, any event downstream of uAUG1 that slows down the scanning 40S subunits upstream of uAUG1 results in a more careful scanning process and a better recognition of the initiation codon. A comparable mechanism was described for Fli-1 (13). Fli-1 translation can be initiated at two in-frame AUG codons, resulting in the synthesis of two protein isoforms of 51 and 48 kDa. The ratio between these isoforms was determined by two uORFs, both overlapping with the 51 kDa Fli-1 initiation codon. Translation of the uORFs stimulated initiation at both Fli-1 AUG codons. Synthesis of the 48 kDa isoform was dependent on reinitiation after translation of a uORF, whereas synthesis of the 51 kDa form was stimulated by scanning 40S subunits queued upstream of terminating ribosomes (13).

It can be argued that the removal of the termination codon of uORF1 caused faster scanning, thereby decreasing the efficiency of uAUG1 recognition, resulting in an underestimation of the initiation frequency at uAUG1 in the #1 transcript. If removal of the uORF1 termination codon indeed caused a poorer recognition of uAUG1, a substantial increase in GFP translation would be expected. This was not found: GFP translation only increased 3-fold. This cannot account for
the 100-fold difference in GFP translation efficiency between the WT and Δ1 Cx41 5’ UTR (Fig. 3).

Inefficient translation due to selective codon usage was shown for the HIV-1 env (envelope glycoprotein) mRNA (20). Conversion of the codon usage of Th1, an abundant cell-surface protein of rodent thymocytes, by the HIV-1 env codon usage severely attenuated translation. In contrast, the replacement of the endogenous codons of GFP, a jellyfish protein relatively poorly translated in mammalian cells, with those of highly expressed mammalian proteins caused an increase in translation. Therefore, codon bias can play a role in the efficiency of translation of mammalian genes (20).

Analyses of codon usage in Caenorhabditis elegans, Drosophila melanogaster and Arabidopsis thaliana showed a strong correlation between codons frequently used to encode a specific amino acid and codons mainly present in abundant transcripts (25) and proteins (26). This correlation is stronger in short than in long proteins (25), which implies that codon usage of short uORFs can play a distinct role in the control of translation efficiency.

Results of the reticulocyte assays did not reflect the results obtained with Xenopus embryos. As shown previously (14), mutation of the uAUGs in the Cx41 5’ UTR had no consistent effect on GFP translation in reticulocyte lysate. This suggested that the uAUGs were inefficiently recognised in vitro. Nevertheless, in vitro translation of the #1 and Δ2Δ3#1 transcripts showed an initiation efficiency at uAUG1 almost identical to the initiation at the GFP AUG. The lack of competition for translation initiation factors in the reticulocyte lysate might induce a better upstream initiation than measured in vivo. For the initiation at uAUG1, but not for uAUG2 and uAUG3, the in vivo circumstances were approached by omitting extra tRNA. With extra tRNA, initiation at uAUG1 and GFP AUG was almost identical, 48% of the ribosomes starting at uAUG1, and this was reduced to 26% without addition of tRNA (Fig. 6). This percentage approaches the in vivo initiation efficiency at uAUG1 (8%, Fig. 7). Further depletion of the reticulocyte lysate from tRNA might further decrease the difference between in vitro and in vivo translation analyses. This suggests that the ratio between mRNA and tRNA or eIFs is important for the translational control of Cx41.

In conclusion, the first uAUG in the Cx41 5’ UTR was recognised with low efficiency, but the presence of this uAUG strongly reduced GFP synthesis. Translation of uORF1 itself was inhibitory, but the effect was stronger due to the rare leucine codon and the termination codon. Remarkably, uAUG2 and uAUG3, the rare leucine codon and the termination codon all induced a more frequent initiation at uAUG1, most probably because the slower scanning induced a better recognition of AUG codons. This suggests that any translational event will cause better scanning of upstream sequences and thus that once a uORF is translated, the queuing 40S subunits will have a higher efficiency of initiation.

Acknowledgements
We thank P. A. Krieg (University of Texas, USA) for the pT7TS clone, C. Kuhlemeier (Berne University, Switzerland) for the eIF4A antibody, the Hubrecht Laboratory (Utrecht, The Netherlands) for providing X. laevis embryos, H. A. Wagemaker for technical assistance, the Department of Image Processing and Design for help in artwork, M. van Hattum for typing the manuscript, and W. J. A. G. Dictus and H. O. Voorma for their interest in our work and for critical reading of the manuscript (all Utrecht University, The Netherlands).

REFERENCES


