A possible contribution of mRNA secondary structure to translation initiation efficiency in *Lactococcus lactis*

Maarten van de Guchte, Ted van der Lende, Jan Kok and Gerard Venema

*Department of Genetics, Centre of Biological Sciences, University of Groningen, Haren, The Netherlands*

Received 15 March 1991
Accepted 25 March 1991

Key words: Gene expression; Translation initiation; mRNA secondary structure; *Lactococcus lactis*

1. SUMMARY

Gene expression signals derived from *Lactococcus lactis* were linked to lacZ-fused genes with different 5′-nucleotide sequences. Computer predictions of mRNA secondary structure were combined with lacZ expression studies to direct base-substitutions that could possibly influence gene expression. Mutations were made such that the DNA sequence upstream of the ATG start codon was not changed. Moreover, care was taken that the substitutions, which were all within the first six codons, neither affected the amino acid sequence of the gene product nor introduced codons rarely used in *L. lactis*. The results suggest that mRNA secondary structure contributes to the efficiency of translation initiation in *L. lactis*.

2. INTRODUCTION

Recently, we described the development of a lactococcal expression vector [1] and its use in the expression of the mature hen egg-white lysozyme (van de Guchte, van der Wal, Kok and Venema, submitted) and of a 5′-truncated *Escherichia coli* lacZ gene [3] in *Lactococcus lactis*. The expression vector, pMG36e, contains a promoter (P32) derived from the chromosome of *L. lactis*, followed by a short open reading frame that encompasses a multiple cloning site (ORF32). In studies on the expression of both hen egg-white lysozyme and β-galactosidase, gene expression was shown to be influenced by the presence of ORF32. The expression levels observed with in-frame fusions to ORF32 were reduced after removal of ORF32. Although this reduction may result from differences in product stability, some evidence was obtained that this reduction was caused by a reduced efficiency of translation initiation [3]. This was inferred from the observation that the expression of the lacZ gene could be enhanced by translational coupling to ORF32. A number of
experimental results prompted us to adopt the translational restart hypothesis proposed by Oppenheim and Yanofsky [4] to explain this coupling. In the context of this hypothesis, we proposed that enhanced lacZ expression was dependent on a higher efficiency of translation initiation of ORF32 than of lacZ. Since the sequences present upstream of the translational start codons of ORF32, of the unfused lacZ gene and of the mature hen egg-white lysozyme coding sequence were identical, sequences within the coding regions must be held responsible for the observed differences in the expression levels.

Coding sequences can affect translation initiation via their contribution to mRNA secondary structure formation. mRNA secondary structures can severely influence gene expression, as has recently been reviewed by de Smit and van Duin [5], and this effect is at the basis of many naturally occurring translational regulation mechanisms, as for example that of the Staphylococcus aureus ermC gene [6]. In the study presented here, we examined whether an influence of mRNA secondary structure could explain the differential expression of hen egg-white lysozyme and of E. coli β-galactosidase observed in the presence or absence of ORF32. We constructed lacZ fusions to corroborate our earlier observations on differential gene expression [2,3], and used computer-predicted secondary structure information to introduce base-substitutions within coding sequences that might influence translation initiation efficiency. The results of this study show that the secondary structure predictions could be successfully used to direct base-substitutions that enhanced gene expression.

3. MATERIALS AND METHODS

3.1. Bacterial strains and media

The bacterial strains used in this study are listed in Table 1. E. coli was grown in TY broth [7] or on TY broth solidified with 1.5% agar. L. lactis was grown in glucose M17 broth [8], or on glucose M17 broth solidified with 1.5% agar. Sucrose (0.3 M) was added to these media to osmotically stabilize electroporated L. lactis cells. Erythromycin was added to a concentration of 100 μg/ml or 5 μg/ml for E. coli and L. lactis, respectively. The chromogenic β-galactosidase substrate 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) (Sigma, St. Louis, MO) was added to a final concentration of 40 μg/ml.

<table>
<thead>
<tr>
<th>Bacterial strain or plasmid</th>
<th>Relevant feature(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacteria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td>araD139, Δ lacZX74, Δ (ara, leu)7697, galU, galK, strA</td>
<td>[29]</td>
</tr>
<tr>
<td>MC1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. lactis</em> subsp. lactis</td>
<td>Plasmid-free strain</td>
<td>[30]</td>
</tr>
<tr>
<td>IL1403</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMG57</td>
<td>Em', lacZ fusion vector</td>
<td>[3]</td>
</tr>
<tr>
<td>pTr1</td>
<td>Em', pMG57 derivative carrying an in-frame fusion of ORF32 and lacZ</td>
<td>This work</td>
</tr>
<tr>
<td>pTr3</td>
<td>Em', pMG57 derivative carrying an in-frame fusion of mature HEL cDNA and lacZ</td>
<td>This work</td>
</tr>
<tr>
<td>pTr3m</td>
<td>Em', pTr3 derivative</td>
<td>This work</td>
</tr>
<tr>
<td>pMG60</td>
<td>Em', carrying the <em>E. coli</em> lacZ gene, in which the first 8 codons were replaced by 3 others</td>
<td>[3]</td>
</tr>
<tr>
<td>pMG60m</td>
<td>Em', pMG60 derivative</td>
<td>This work</td>
</tr>
</tbody>
</table>
3.2. Plasmid constructions

The plasmids used in this study are listed in Table 1. Plasmids containing lacZ gene fusions are all based on pMG57 [3] (Fig. 1). Plasmid pMG57 contains the broad-host-range origin of replication derived from pWV01 [9], which enables the plasmid to replicate in E. coli, L. lactis, and several other bacterial species. In this plasmid, a lactococcal promoter (P32), a short open reading frame (ORF32) and a unique SmaI restriction site are present upstream of a 5'-truncated E. coli lacZ gene. P32 and the first nine codons of ORF32 were isolated as one unit from the chromosome of L. lactis subsp. cremoris Wg2 [10]. The lacZ gene in pMG57 has a reading frame different from that of ORF32. An in-frame fusion of the two open reading frames was obtained by the insertion of two nucleotides at the SmaI site in pMG57. To this purpose, pMG57 was cut with XmaI, after which the 5'-protruding ends were filled in with Klenow enzyme. The plasmid was subsequently cut with HindIII, and the lacZ fragment was isolated. A second batch of pMG57 was cut with SmaI and HindIII, after which the part of the plasmid containing the erythromycin resistance marker was isolated and ligated to the lacZ fragment to give pTr1. In the fused gene, the ninth codon of E. coli lacZ is preceded by the 48 nucleotides of ORF32.

In pTr3, lacZ is preceded by the first part of the mature hen egg-white lysozyme (HEL) coding sequence derived from pMG37 [2]. Expression signals identical to those present upstream of ORF32 in pMG57 are present upstream of the mature HEL encoding sequence in pMG37. A HindII site was created within the HEL coding sequence by site-directed mutagenesis. The EcoRI–HindII fragment containing P32 and the first part of the HEL sequence was subsequently used to replace the corresponding EcoRI–SmaI fragment of pMG57. Thus, in pTr3 the ninth

Fig. 2. 5'-terminal mRNA sequences produced from the lacZ-fusions present on the plasmids indicated. *, starting point of transcription [10]; AUG, translational start codon; Shine Dalgarno sequences are doubly underlined; differences between the sequences of pTr3 and pTr3m, and between pMG60 and pMG60m are indicated by dashed lines. Numbers indicate positions from the start of the messenger. See text for details on coding sequences.
Fig. 3. Computer-predicted secondary structures of the 5'-terminal mRNA sequences specified by the plasmids indicated and shown in Fig. 2. For the mRNA from pTr3, two alternative structures are given. Shine Dalgarno sequences and AUG translational start codons are boxed. Arrows indicate the sites at which base-substitutions were introduced in pTr3 and pMG60. Centred dots indicate juxtaposed bases that are not Watson-Crick base-pairs.
codon of *E. coli* lacZ is preceded by a sequence of 48 nucleotides consisting of an ATG start codon, 43 nucleotides of the mature HEL coding sequence, and two linking nucleotides.

Plasmid pMG60 was described previously [3]. In this plasmid, the ninth codon of *lacZ* is preceded by an ATG start codon, and six linking nucleotides. In both pTr3 and pMG60, three nucleotides within the coding sequence were changed by site-directed mutagenesis to give pTr3m and pMG60m, respectively (Fig. 2). All lacZ fusions were constructed in *E. coli* MC1000, with selection for erythromycin resistance and blue colour development on TY agar containing X-gal. Fusion points and promoter areas were checked by DNA sequencing.

3.3. DNA manipulations and transformation

Plasmid DNA was isolated by the method of Birnboim and Doly [11]. Restriction enzymes, Klenow enzyme, and T4 DNA ligase were purchased from Boehringer (Mannheim, F.R.G.) and used according to the instructions of the supplier. Site-directed mutagenesis was performed using the PMa/c gapped duplex method described by Stanssens et al. [12]. Oligonucleotides were synthesized using an Applied Biosystems 381A DNA synthesizer (Applied Biosystems, Inc., Foster City, CA) and had the following 5′→3′ sequences: CCG ATA GTT ATC AAG TCG TCG ACG CTT CAT A (used in the construction of pTr3); TGC CAG CTC ACA CCT TCC AAA CAC TTT CAT (used in the construction of pTr3m); ACG ACG TTG TAA CAC CAC TGC CCC TTT CAT (used in the construction of pMG60m). DNA sequence analysis was performed according to Tabor and Richardson [13] using the T7 polymerase system (Pharmacia, Uppsala, Sweden) on denatured plasmid DNA. *E. coli* was transformed according to the protocol of Mandel and Higa [14]. Plasmids were introduced into *L. lactis* by electroporation [15].

3.4. Secondary structure analysis

5′-Terminal mRNA sequences including the first 25 nucleotides of the coding sequences (Fig. 2) were analyzed using the computer programs Fold [16,17] and ASRNA [18]. Both programs use the free-energy parameters of Freier et al. [19].

3.5. Assay of β-galactosidase activity

β-Galactosidase activity per mg protein was determined in lysates of overnight-grown cultures of *L. lactis*, essentially by the method of Miller [20] as described previously [3]. The protein content of the lysates was determined according to Bradford [21].

4. RESULTS

4.1. Secondary structure analysis

To examine the relationship between the level of gene expression and the presence of ORF32, potential secondary structures in the 5′-terminal mRNA sequences of the lacZ fusions in pTr1, pTr3, and pMG60 were determined. The results presented in Fig. 3 show that in the structure predicted for pTr1 the Shine Dalgarno (SD) region is not involved in secondary structure formation, whereas this region may be occluded in pTr3 and pMG60. To test the effect of the proposed secondary structures on lacZ expression, base-substitutions were introduced within the coding sequences of the lacZ fusions in pTr3 and pMG60 with the aim of destabilizing these structures (Fig. 2). The base-substitutions were chosen such that neither the amino acid sequence of the gene product would be affected, nor would a significant effect of codon usage on the expression in *L. lactis* be expected. The 5′-terminal mRNA sequences of the lacZ fusions in the resulting plasmids pTr3m and pMG60m were also subjected to computer-folding. The results presented in Fig. 3 show that the mutations introduced into pTr3 to produce pTr3m diminished the overall free energy of the proposed secondary structure that comprises the SD region. Concomitantly the AUG start-codon no longer participated in base-pairing. The SD region in the mRNA derived from the lacZ-fusion in pMG60m remained occluded in a putative hairpin that was also present in the messenger from pMG60. The secondary structure in which the translational
start codon was involved in the mRNA from pMG60 had disappeared, however.

4.2. Gene expression in \textit{L. lactis}

To determine the effect of the introduced base-substitutions on gene expression, \(\beta\)-galactosidase activity was measured in lysates of \textit{L. lactis} strains harbouring the different plasmids. The results of this assay are presented in Fig. 4. \textit{L. lactis} IL1403(pTr1), carrying an ORF32-\textit{lacZ} fusion, showed an activity level significantly higher than that of IL1403(pTr3), in which the 5'-truncated \textit{lacZ} was fused to a short hen egg-white lysozyme sequence. The \(\beta\)-galactosidase activity displayed by strain IL1403(pTr1) also far exceeded that of IL1403(pMG60), in which the \textit{lacZ} sequence was preceded by an ATG and two linking codons. In strain IL1403(pTr3m) an activity level 3.6 times higher than that of IL1403(pTr3) was observed. The expression levels observed in IL1403(pMG60) and IL1403(pMG60m) were similar. These results show that minor changes in a coding sequence can highly affect gene expression.

5. DISCUSSION

The secondary structure of mRNA can severely affect gene expression by influencing translation initiation efficiency [5,22–26]. In the study presented here, we examined whether mRNA secondary structure could account for differential expression of \textit{lacZ} fusions in \textit{L. lactis}. A 5'-truncated \textit{E. coli lacZ} gene was fused to ORF32 present in pMG57(pTr1) or to the 5'-part of the mature hen egg-white lysozyme (HEL) coding sequence (pTr3). In pMG60, the \textit{lacZ} sequence is preceded by a sequence of three codons. Computer folding of the first 55 nucleotides of the various mRNAs indicated that the Shine Dalgarno (SD) regions present in pTr3 and pMG60 may be occluded in secondary structures, whereas in pTr1 this sequence seemed to be freely accessible. Determination of \(\beta\)-galactosidase activity in \textit{L. lactis} strains harbouring the various plasmids, showed that the activity in IL1403(pTr1) far exceeded that in IL1403(pTr3) and IL1403(pMG60). Although the plasmids specified \(\beta\)-galactosidases with different N-terminal extensions, which may differ in stability or specific activity, a causal connection may exist between the accessibility of the SD sequence and the observed \(\beta\)-galactosidase activity. In order to test this hypothesis, base-substitutions were introduced in pTr3 and pMG60 to destabilize the proposed secondary structures. The substitutions were introduced only in the coding sequences, since our main interest was to study how different coding sequences combine with the specific expression signals located upstream of ORF32 in pMG57 and in the lactococcal expression vector pMG36e. In addition, we limited ourselves to substitutions that would not affect the amino acid sequence of the gene product and care was taken not to introduce codons that are rarely used in \textit{L. lactis} (based on the codon usage in the \textit{L. lactis} subsp. cremoris Wg2 proteinase gene [27]). These changes resulted in a significant reduction of the overall free energies of the proposed secondary structures in the mRNAs. In pTr3m this decrease in secondary structure might result in an improved accessibility of the ribosome binding site. In pMG60, however, the hairpin that comprised the SD region, could not be removed under the constraint not to change any of the nucleotides upstream of the coding sequence. The additional stem-loop structures in the mRNA specified by pMG60, in which the coding sequence participated, had disappeared in pMG60m. Concomitantly, the AUG start codon became exposed in pMG60m, whereas
it was occluded in pMG60. In IL1403, pTr3m gave rise to a β-galactosidase activity level 3.6 times higher than that of IL1403(pTr3), suggesting that mRNA secondary structure indeed contributed to translation initiation efficiency in this case. The β-galactosidase activities observed in IL1403(pMG60) and IL1403(pMG60m) were similar, suggesting that especially the secondary structure in which the SD region is involved is important in determining the efficiency of translation initiation. This observation would be in accordance with the suggestion made by Hager and Rabinowitz [28], that in Gram-positives an exposed SD sequence may be more important than an exposed start codon for an efficient initiation of translation.

In both pTr3 and pMG60, we were able to remove putative mRNA secondary structures involving the coding sequences by base-substitutions in these sequences. The remaining secondary structures only involved sequences upstream of the AUG start codon. The important feature that distinguishes pTr1 from pTr3m and pMG60m is that in the former the coding sequence can base-pair with the AUG start codon and the sequence directly preceding it. This base-pairing prevents the pairing of the sequence directly upstream of the AUG with the SD region in pTr1, a pairing that does occur in the mRNAs specified by pTr3m and pMG60m.

If the mRNA secondary structures predicted by computer analysis also exist in vivo, and if the mRNA level is not affected by the changes introduced, the results presented here strongly suggest that mRNA secondary structure is an important determinant of translation initiation efficiency in the P32-ORF32 configuration in L. lactis. As a consequence, replacement of ORF32 by another coding sequence can easily interfere with an efficient initiation of translation. Therefore, a configuration of translational coupling may be preferred over a precise fusion of homologous expression signals and a heterologous coding sequence to obtain efficient expression of the heterologous gene. The strategy of coupling the translation of a heterologous gene to that of a homologous upstream situated coding sequence has already been successfully applied to improve the expression of the E. coli lacZ gene in L. lactis [3].

ACKNOWLEDGEMENTS

This work was supported by BCZ Friesland, Leeuwarden, The Netherlands. We thank Henk Mulder for preparation of the figures.

REFERENCES