

The increased level of β 1,4-galactosyltransferase required for lactose biosynthesis is achieved in part by translational control

(5'-untranslated region/*in vitro* translation/*in vivo* translation/mammary gland/RNA secondary structure)

MARTIN CHARRON*, JOEL H. SHAPER*[†], AND NANCY L. SHAPER*[‡]

*Cell Structure and Function Laboratory and Oncology Center and [†]Department of Pharmacology and Molecular Sciences, School of Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD 21287-8937

Communicated by Robert L. Hill, Duke University Medical Center, Durham, NC, October 5, 1998 (received for review March 27, 1998)

ABSTRACT β 1,4-Galactosyltransferase (β 4GalT-I) participates in both glycoconjugate biosynthesis (ubiquitous activity) and lactose biosynthesis (mammary gland-specific activity). In somatic tissues, transcription of the mammalian β 4GalT-I gene results in a 4.1-kb mRNA and a 3.9-kb mRNA as a consequence of initiation at two start sites separated by \approx 200 bp. In the mammary gland, coincident with the increased β 4GalT-I enzyme level (\approx 50-fold) required for lactose biosynthesis, there is a switch from the 4.1-kb start site to the preferential use of the 3.9-kb start site, which is governed by a stronger tissue-restricted promoter. The use of the 3.9-kb start site results in a β 4GalT-I transcript in which the 5'-untranslated region (UTR) has been truncated from \approx 175 nt to \approx 28 nt. The 5'-UTR of the 4.1-kb transcript [UTR(4.1)] is predicted to contain extensive secondary structure, a feature previously shown to reduce translational efficiency of an mRNA. In contrast, the 5'-UTR of the 3.9-kb mRNA [UTR(3.9)] lacks extensive secondary structure; thus, this transcript is predicted to be more efficiently translated relative to the 4.1-kb mRNA. To test this prediction, constructs were assembled in which the respective 5'-UTRs were fused to the luciferase-coding sequence and enzyme levels were determined after translation *in vitro* and *in vivo*. The luciferase mRNA containing the truncated UTR(3.9) was translated more efficiently both *in vitro* (\approx 14-fold) and *in vivo* (3- to 5-fold) relative to the luciferase mRNA containing the UTR(4.1). Consequently, in addition to control at the transcriptional level, β 4GalT-I enzyme levels are further augmented in the lactating mammary gland as a result of translational control.

β 1,4-Galactosyltransferase (β 4GalT-I) is a constitutively expressed, trans-Golgi resident, type II membrane-bound glycoprotein that is widely distributed in vertebrates. β 4GalT-I catalyzes the transfer of galactose to *N*-acetylglucosamine residues, forming the β 4-*N*-acetyllactosamine (Gal β 4GlcNAc) or poly β 4-*N*-acetyllactosamine structures found in glycoconjugates (1). In mammals, β 4GalT-I has been recruited for a second biosynthetic function, the tissue-specific production of lactose (Gal β 4Glc), which takes place exclusively in the lactating mammary gland. The synthesis of lactose is carried out by a protein heterodimer assembled from β 4GalT-I and α -lactalbumin, a noncatalytic mammalian protein expressed *de novo* exclusively in the mammary gland during lactation (2–4). The net result of the association of α -lactalbumin with β 4GalT-I is to lower the K_m for glucose about three orders of magnitude, consequently making it an effective acceptor substrate at physiological concentrations.

Beginning in late pregnancy, the β 4GalT-I enzyme level in the mammary gland has been estimated to increase \approx 50-fold in preparation for lactose biosynthesis (5, 6). With the recruitment of β 4GalT-I for lactose biosynthesis, the regulatory problem arose as to how to specifically increase the enzyme level only in the lactating mammary gland while maintaining the comparatively low level of constitutively expressed enzyme required for glycoconjugate biosynthesis in all other somatic tissues. To address this question, we have carried out a detailed analysis of the structure and transcriptional regulation of the murine β 4GalT-I gene. The main observations can be summarized as follows. (i) The murine β 4GalT-I gene specifies two transcripts of \approx 4.1 and \approx 3.9 kb in somatic cells as a consequence of initiation at two different start sites separated by \approx 200 bp (ref. 7; Fig. 1A). The identical structural features are also found in the bovine (8) and human (9, 10) β 4GalT-I genes, which suggests that they may be a distinguishing characteristic of all mammalian β 4GalT-I orthologues. (ii) Structurally, the 4.1- and 3.9-kb transcripts are essentially identical, with the exception of their respective 5'-untranslated region (UTR). The length of the 5'-UTR of the 4.1-kb transcript [UTR(4.1)] is \approx 175 nt; in contrast, the 5'-UTR of the 3.9-kb transcript [UTR(3.9)] is only \approx 28 nt long (Fig. 1B). (iii) In murine somatic tissues, the 4.1-kb start site is used predominantly. Expression from this start site is controlled by a relatively weak promoter containing multiple Sp1 sites, consistent with the role of the 4.1-kb mRNA as the ubiquitous transcript involved in glycoconjugate biosynthesis (11, 12). (iv) In the mid- to late-pregnant and lactating mammary gland, the 3.9-kb start site is used preferentially. Expression from this start site is controlled by a stronger promoter that is in part regulated by lactating mammary gland-restricted transcription factors. The net result of this switch to the 3.9-kb start site in the lactating mammary gland is a \approx 10-fold increase in the steady-state levels of β 4GalT-I mRNA (11, 12).

In the context of the required \approx 50-fold increase in β 4GalT-I enzymatic activity in the lactating mammary gland, two questions arose. First, could the \approx 10-fold increase in mRNA levels by itself account for the \approx 50-fold increase in enzyme levels observed? Second, why has nature gone to the trouble to generate a second β 4GalT-I transcript that is distinguished by the absence of \approx 180 nt, primarily from the 5'-UTR? A comparison of the UTR(4.1) and UTR(3.9) provided the

Abbreviations: β 4GalT-I, β 1,4-galactosyltransferase (EC 2.4.1.38); murine β 4GalT-I refers to the α -lactalbumin-responsive UDP galactose, *N*-acetyl- β -D-glucosaminylglycopeptide β 1,4-galactosyltransferase that has been mapped to the centromeric region of mouse chromosome 4; GlcNAc, *N*-acetylglucosamine; UTR, untranslated region; UTR(4.1), 5'-UTR of the 4.1-kb β 4GalT-I mRNA; UTR(3.9), 5'-UTR of the 3.9-kb β 4GalT-I mRNA; LUC, luciferase; eIF, eukaryotic initiation factor; RT-PCR, reverse transcription-PCR.

[‡]To whom reprint requests should be addressed at: The Johns Hopkins University School of Medicine, Oncology Center, Room 1-127, 600 North Wolfe Street, Baltimore, MD 21287-8937. e-mail: nshaper@jhmi.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/9514805-6\$2.00/0 PNAS is available online at www.pnas.org.

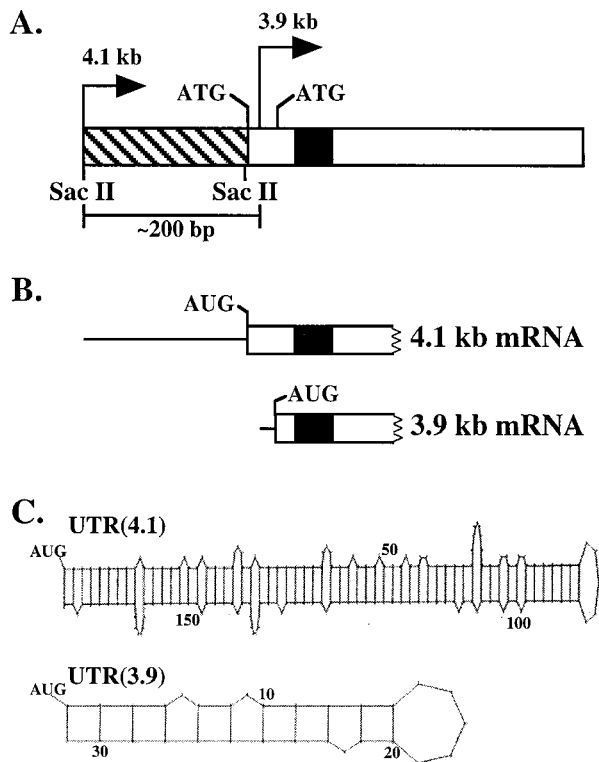


FIG. 1. (A) Schematic representation of exon one of the murine $\beta 4$ GalT-I gene. The bent arrows denote the position of the 4.1-kb transcriptional start site (4.1 kb) and the 3.9-kb transcriptional start site (3.9 kb) relative to the two in-frame ATGs that are separated by 39 bp. The 3.9-kb start site is positioned between the two in-frame ATGs, ≈ 200 bp downstream of the 4.1-kb start site. The hatched and open rectangles represent the 5'-UTR and coding sequence of the 4.1-kb mRNA, respectively. Note the UTR(3.9) is embedded within the coding sequence of the 4.1 kb-mRNA. The black box indicates the position of the transmembrane domain. (B) The 5' end of the 4.1- and 3.9-kb $\beta 4$ GalT-I mRNA is shown, where the thin line and open rectangle represent the 5'-UTR and part of the coding sequence, respectively. Translation of the 4.1- and 3.9-kb $\beta 4$ GalT-I mRNA results in two catalytically identical, trans-Golgi resident protein isoforms with NH₂-terminal cytoplasmic domains of 24 and 11 aa, respectively. (C) Potential secondary structure of the 5'-UTR of the 4.1- and 3.9-kb $\beta 4$ GalT-I transcript. The long UTR(4.1) has a calculated ΔG of -76 kcal/mol, whereas the short UTR(3.9) has a calculated ΔG of only -7 kcal/mol (13).

necessary clue. The long UTR(4.1) has the potential to form an almost perfect hairpin structure with a calculated free energy of formation (ΔG) of -76 kcal/mol (Fig. 1C; ref. 13). It has been shown that 5'-UTRs with ΔG s < -50 kcal/mol significantly reduce translational efficiency (ref. 14, and reviewed in ref. 15). In contrast, the calculated ΔG of the short UTR(3.9) is only -7 kcal/mol (Fig. 1C). Consequently, the 3.9-kb $\beta 4$ GalT-I transcript is predicted to be more efficiently translated relative to the 4.1-kb transcript.

The goal of this study was to experimentally test this prediction. Luciferase (LUC) constructs in which the respective 5'-UTRs were inserted between the SP6 or simian virus (SV) 40 transcriptional start site and the initiation codon of LUC were assembled. By using a cell-free system or transfection into cell lines, we show that the LUC mRNA with the UTR(3.9) is translated more efficiently compared with the mRNA with the UTR(4.1). Consequently, in addition to control at the transcriptional level, $\beta 4$ GalT-I enzyme levels in the lactating mammary gland are further augmented as a result of translational control.

MATERIALS AND METHODS

LUC Constructs. All constructs were derived from the pGL2-promoter expression vector (Promega) in which the

SV40 transcriptional start site is 91 bp upstream of the LUC initiation codon (Fig. 2A). pGL2 was digested with *Sfi*I, which cleaves precisely at the SV40 transcriptional start site (nucleotide 176) and *Nar*I, which cleaves 34 bp downstream of the ATG initiation codon. A double-stranded oligomer containing a *Sfi*I overhang at the 5' end, the 28-bp sequence of the UTR(3.9) plus the first 34 bp of the LUC coding sequence, and a *Nar*I overhang at the 3' end was ligated into the *Sfi*I-*Nar*I site, yielding the construct SV40-UTR(3.9)-LUC (Fig. 2B). SV40-UTR(4.1)-LUC was constructed essentially as above by using a double-stranded synthetic oligomer containing an *Sfi*I overhang at the 5' end, the *Sac*II site, the last 7 bp of the 175-bp UTR(4.1) sequence, the first 34 bp of the LUC coding sequence, and a *Nar*I overhang at the 3' end. The resulting construct was digested with *Sac*II to insert the *Sac*II fragment containing the initial 162 bp of the UTR(4.1) sequence (Fig. 2C). To assemble the SP6-UTR(3.9)-LUC construct (Fig. 2B), SV40-UTR(3.9)-LUC was digested with *Bgl*II, which cleaves pGL2 at nucleotide 37, and *Alw*NI, which cleaves pGL2 at nucleotide 3,602 and also cleaves inside the UTR(3.9) sequence, thus removing the SV40 promoter and the first 10 bp of the UTR(3.9). A double-stranded oligomer was synthesized that contained a *Bgl*II overhang at the 5' end, the SP6 promoter sequence 5'-ATTTAGGTGACACTATA-3' (16), and the first 10-bp sequence of the UTR(3.9) containing an *Alw*NI overhang at the 3' end. A three-way ligation between the *Alw*NI fragment (containing the last 18 bp of the UTR(3.9), the LUC coding sequence, and the SV40 3'-UTR), the *Alw*NI-*Bgl*II fragment (containing nt 3,602-37 of pGL2), and the oligomer yielded the construct SP6-UTR(3.9)-LUC. The first T of the UTR(3.9) sequence was changed to a G; this transversion results in a GGA triplet and leads to increased transcriptional activity from the SP6 promoter (16). Finally, to assemble the SP6-UTR(4.1)-LUC construct (Fig. 2C), SV40-UTR(4.1)-LUC was digested with *Bgl*II and *Sac*II, removing the SV40 promoter and the initial 166 bp of the UTR(4.1). A double-stranded oligomer containing a *Bgl*II overhang at the 5' end, the SP6 promoter, the GGA triplet, and a *Sac*II overhang at the 3' end was ligated into the *Bgl*II-*Sac*II site. The resulting plasmid was digested with *Sac*II to insert the *Sac*II fragment containing the initial 162 bp of the UTR(4.1) sequence, yielding SP6-UTR(4.1)-LUC. Sequencing confirmed the structure of each construct.

Cells and Culture Conditions. The monkey COS-1 and human embryonal kidney 293 cell lines were cultured in DMEM with high glucose that contained 10% fetal bovine serum, 100 units/ml penicillin, and 50 μ g/ml streptomycin. Mouse L cells (American Type Culture Collection) were cultured as described (12). Media was from Life Technologies (Grand Island, NY).

Transfection and Reporter Gene Assays. COS-1 cells were transfected as recommended by Peter Barry (University of California, Davis CA, personal communication). COS-1 cells ($\approx 75\%$ confluent) were trypsinized, centrifuged at $500 \times g$ for 10 min, washed twice with PBS, and resuspended in RPMI medium 1640 without serum at $\approx 1.3 \times 10^7$ cells per ml. Aliquots (0.3 ml) of cells plus 10 μ g of either SV40-UTR(4.1)-LUC or SV40-UTR(3.9)-LUC were electroporated using 960 μ F at 250 V (Bio-Rad gene pulser). Under these conditions, the transfection efficiency is $>80\%$. Mouse L cells ($\approx 75\%$ confluent) were collected as above, washed twice in ice-cold OptiMEM, and resuspended in the same medium at $\approx 2.5 \times 10^7$ cells per ml. Aliquots (0.4 ml) of cells plus 10 μ g of each construct were placed in electroporation cuvettes and incubated for 10 min on ice before electroporation using 960 μ F at 300 V. After electroporation, samples were kept at room temperature for 10 min and then transferred into 10 ml of medium; the transfection efficiency was $\approx 2\%$. Human 293 cells were transfected with 100 ng of each construct by using the calcium phosphate method (17); the transfection efficiency

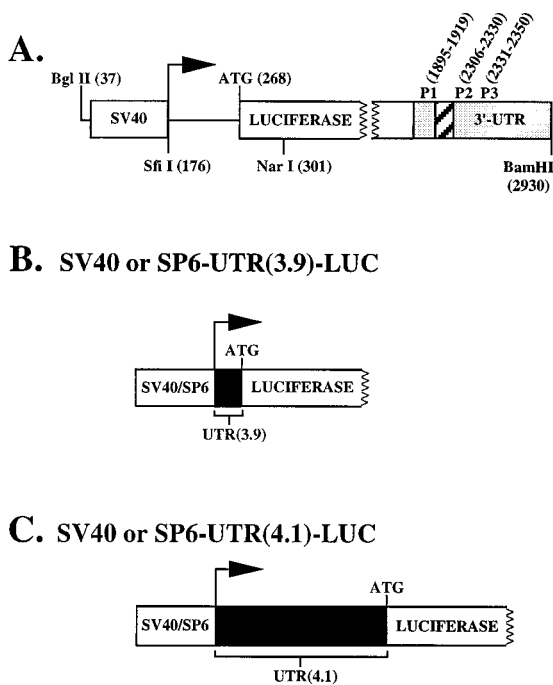


FIG. 2. (A) Schematic representation of the pGL2-promoter vector. The horizontal arrow indicates the SV40 transcriptional start site, the open box designated SV40 indicates the SV40 promoter, the open box designated LUC indicates the LUC coding sequence, the shaded box indicates the SV40 3'-UTR, and the hatched box indicates the position of the intron. P1 (forward), P2 (reverse), and P3 (reverse) designate the position of the primers used for the RT-PCR experiments. Numbers in parentheses indicate relevant nucleotide positions discussed in the text. (B) Schematic representation of the LUC construct containing either the SV40 or SP6 promoter plus UTR(3.9). (C) Schematic representation of the LUC construct containing either the SV40 or SP6 promoter plus UTR(4.1).

was $\approx 80\%$. After transfection, cells were incubated at 37°C in a humidified atmosphere of $5\% \text{CO}_2$ for 24 or 48 hr, at which time they were assayed for LUC activity as described (18). To normalize transfection efficiency, cells were cotransfected with the β -galactosidase plasmid, pON260 (11), and activity was measured by using a chemiluminescence assay (Galactolight; Tropix, Bedford, MA) according to manufacturer's instructions. Chemiluminescent emission was measured for 10 sec by using a Monolight 500 luminometer (Analytical Luminescence Laboratory, San Diego).

RNA Isolation, Reverse Transcription-PCR (RT-PCR), and S1 Nuclease Analysis. After total RNA was isolated from transfected COS-1 cells as described (19), LUC mRNA was quantitated by using a previously described RT-PCR protocol (20) with slight modifications. Briefly, after digestion with 1 unit of amplification-grade DNase I (Life Technologies, Grand Island, NY) for 30 min at 37°C , $1 \mu\text{g}$ of RNA was reverse-transcribed with 50 pmol of primer P3-reverse (nucleotides 2,331–2,350) by using the SuperScript preamplification system according to the manufacturer's instructions (Life Technologies, Grand Island, NY). An aliquot ($2 \mu\text{l}$) of the RT reaction was added to $48 \mu\text{l}$ of a PCR mixture containing $10 \mu\text{l}$ of $5\times$ buffer B (Invitrogen), 2.5 mM dNTPs, $2 \mu\text{Ci}$ ($1 \text{ Ci} = 37 \text{ GBq}$) of $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$, 2.5 units of *Taq* DNA polymerase (Life Technologies), and 50 pmol of primers P1-forward (nucleotides 1,895–1,919) and P2-reverse (nucleotides 2,306–2,330; Fig. 2A). PCR was carried out for 22 cycles at 95°C for 1 min, 68°C for 1 min, and 72°C for 1 min, with a final extension at 72°C for 3 min. The products were electrophoresed on an agarose gel and transferred to Nytran (Schleicher & Schuell). The relative intensities of the bands were measured by using a

PhosphorImager (Molecular Dynamics). S1 nuclease protection analysis was carried out as described (7). Two single-stranded probes were generated to distinguish between the UTR(4.1)–LUC and UTR(3.9)–LUC mRNA: one probe (341 nt in length) contained the last 46 bp of the SV40 promoter sequence, the UTR(4.1) sequence, and the initial 120 bp of the LUC-coding sequence, whereas the other (401 nt in length) contained the last 46 bp of the SV40 promoter sequence, the UTR(3.9) sequence, and the initial 327 bp of the LUC-coding sequence.

In Vitro Transcription and Translation. *In vitro* transcription was performed with the SP6 Riboprobe System from Promega according to manufacturer's instructions using $2 \mu\text{g}$ of plasmid SP6–UTR(4.1)–LUC or SP6–UTR(3.9)–LUC linearized with *Bam*HI (Fig. 2A). Following digestion with 2 units of RNase-free DNase (Promega) for 15 min at 37°C , RNA was extracted with phenol/chloroform, precipitated with NH_4Ac /ethanol, washed with 70% ethanol, and quantitated spectrophotometrically. Integrity of the RNA was determined by formaldehyde-agarose gel electrophoresis. In some experiments, capped RNA was prepared by using guanylyltransferase (Life Technologies, Grand Island, NY) according to manufacturer's instructions (Promega). *In vitro* translation was performed by using a rabbit reticulocyte lysate system (Promega) according to manufacturer's instructions. RNA ($0.5 \mu\text{g}$ in $1 \mu\text{l}$) was added to $11 \mu\text{l}$ of H_2O , heated at 70°C for 10 min, quick-cooled on ice, and immediately added to the reticulocyte lysate mixture ($50 \mu\text{l}$) and incubated at 30°C for various times. Aliquots ($5 \mu\text{l}$) were assayed for LUC activity at different time intervals.

RESULTS

Experimental Strategy. The experiments presented were designed to test the relative effect of the 5'-UTR of the 3.9- and 4.1-kb $\beta\text{4GalT-I}$ mRNA on the translational efficiency of LUC mRNA both *in vitro* and *in vivo*. Care was therefore taken to assemble the LUC constructs such that each 5'-UTR was precisely inserted between the SP6 or SV40 transcriptional start site and the ATG initiation codon of LUC. The SP6 and SV40 promoters were chosen because in both cases, the precise nucleotide at which transcription initiation occurs is known (16, 21, 22). Because the 4.1-kb mRNA contains multiple closely spaced transcriptional start sites that range from 160 nt to 190 nt upstream of the first in-frame ATG (7), an average 5'-UTR length of 175 nt was used.

Effect of the UTR(3.9) and UTR(4.1) on Translation *in Vitro*. To determine the effect the 5'-UTR of each $\beta\text{4GalT-I}$ transcript has on translation efficiency *in vitro*, two plasmids were assembled that contained the SP6 promoter sequence and either the UTR(3.9) or the UTR(4.1) fused to the LUC-coding sequence. mRNA was generated from SP6–UTR(3.9)–LUC and SP6–UTR(4.1)–LUC by using SP6 polymerase, and an equal amount ($0.5 \mu\text{g}$) of each mRNA was added to a rabbit reticulocyte lysate system. Translational efficiency was determined by measuring LUC activity as a function of time (0–120 min). The results show that translation of the UTR(3.9)–LUC mRNA consistently gave rise to more LUC activity than translation of the UTR(4.1)–LUC mRNA (Fig. 3A). Data from three independent sets of *in vitro*-transcribed RNA established that translation of UTR(3.9)–LUC mRNA was 14- to 20-fold higher compared with that of UTR(4.1)–LUC mRNA (Fig. 3B).

A comparison was also made between capped and uncapped transcripts because the presence of the m^7G cap structure has been shown to increase the efficiency of translation of some mRNAs *in vitro* (15). Data from two independent sets of *in vitro*-transcribed capped RNA established that the LUC activity from both capped UTR(3.9)–LUC- and UTR(4.1)–LUC mRNA was ≈ 1.5 - to 2-fold higher relative to the correspond-

ing uncapped mRNA. However, the measured translational efficiency of the capped UTR(3.9)-LUC mRNA relative to the capped UTR(4.1)-LUC mRNA remained ≈ 14 -fold higher (data not shown).

Effect of the UTR(3.9) and UTR(4.1) on Translation *in Vivo*.

The experiments described above demonstrate that, *in vitro*, an mRNA containing the UTR(3.9) is translated 14- to 20-fold more efficiently compared with an mRNA containing the UTR(4.1). To establish whether the same effect is observed *in vivo*, constructs were assembled that contained the SV40 promoter sequence fused to either the UTR(3.9) or the UTR(4.1) followed by the LUC coding sequence (Fig. 2). Each construct was transiently transfected into COS-1 cells. S1 nuclease analysis established that transcription initiation occurred at the predicted start site in the SV40 promoter sequence for each of the constructs (data not shown). Translational efficiency was determined by measuring LUC activity (Table 1). When assayed 24 hr after transfection, enzyme activity was ≈ 3 -fold (2.5 ± 0.1 , $n = 3$) higher in cells transfected with SV40-UTR(3.9)-LUC compared with cells transfected with SV40-UTR(4.1)-LUC. A slight increase in

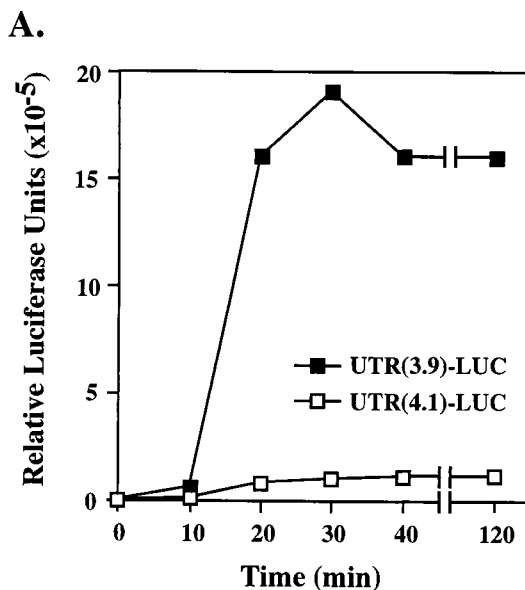
this value (2.8 ± 0.4 , $n = 5$) was observed at 48 hr posttransfection. To show that this effect was not cell-type specific, LUC activity was also measured in mouse L cells and human embryonal kidney 293 cells. Enzyme activity was ≈ 3 -fold higher in L cells and ≈ 5 -fold higher in 293 cells transfected with SV40-UTR(3.9)-LUC compared with cells transfected with SV40-UTR(4.1)-LUC (data not shown).

To establish that the increase in LUC activity reflects increased translation and is not the result of a difference in the level of mRNA transcribed *in vivo*, LUC mRNA was quantitated by using an RT-PCR assay (20). To calibrate the assay method and ensure that a 3-fold difference in mRNA levels could be detected, a standard curve was first generated by using *in vitro*-transcribed LUC mRNA that was reverse-transcribed using primer P3. Primers P1 and P2 were then used for PCR in the presence of [32 P]dCTP and the products were quantitated by PhosphorImager after electrophoresis and transfer to Nytran. As seen in Fig. 4, the assay is linear in the range of 0.4–100 pg of RNA, and a 3-fold difference is readily detected.

Subsequently, total RNA was isolated from transiently transfected COS-1 cells (experiments 4 and 5; Table 1) and reverse-transcribed with primer P3 and subjected to PCR using primers P1 and P2. As seen in Table 2, the levels of LUC mRNA are essentially identical from cells transfected with SV40-UTR(4.1)-LUC or SV40-UTR(3.9)-LUC. Data from three independent RT-PCR experiments established that the RNA ratio between cells transfected with the SV40-UTR(3.9)-LUC construct and cells transfected with the SV40-UTR(4.1)-LUC construct was 1.1 ± 0.2 in experiment 4; a ratio of 0.9 ± 0.2 was obtained by using RNA samples from experiment 5. Therefore, the increased amount of LUC activity detected in cells transfected with the SV40-UTR(3.9)-LUC construct is not the result of higher steady-state levels of this mRNA but is a direct consequence of increased translational efficiency of the UTR(3.9)-LUC mRNA relative to the UTR(4.1)-LUC mRNA.

DISCUSSION

Mammals Have Evolved a Two-Step Mechanism to Generate the Elevated Levels of $\beta 4$ GalT-I Enzymatic Activity Required for Lactose Biosynthesis. The genesis of this study was our proposed model for the transcriptional and translational regulation of the murine $\beta 4$ GalT-I gene, which accounted for the ≈ 50 -fold increase in $\beta 4$ GalT-I enzymatic activity selectively in the mammary gland during lactation (11). We postulated that a two-step mechanism is used to generate the requisite levels of $\beta 4$ GalT-I required for lactose biosynthesis. In step one, steady-state $\beta 4$ GalT-I mRNA levels are up-



B.

Time (min)	UTR(3.9)-LUC / UTR(4.1)-LUC
10	16 \pm 5
20	17 \pm 4
30	17 \pm 4
40	14 \pm 1
120	20 \pm 6

FIG. 3. mRNA containing the UTR(3.9) is translated *in vitro* ≈ 14 -fold more efficiently relative to mRNA containing the UTR(4.1). (A) The two SP6 constructs were linearized with *Bam*HI and transcribed *in vitro* by using SP6 polymerase. An equal amount of each RNA (0.5 μ g) was translated by using the rabbit reticulocyte lysate system as described in *Materials and Methods*. Aliquots (5 μ l) from each reaction were assayed for LUC activity at the indicated times. A representative time course is shown. (B) The ratio between the LUC activity generated by the translation of the UTR(3.9)-LUC mRNA relative to that of the UTR(4.1)-LUC mRNA is shown at different time points. Data were averaged from three independent sets of *in vitro*-transcribed RNA.

Table 1. Effect of the UTR(3.9) and UTR(4.1) on translation *in vivo*

Exp.	UTR(3.9)/ UTR(4.1)*	
	24 hr	48 hr
1	2.4	3.3
2	2.6	2.3
3	2.6	2.6
4	nd	2.8
5	nd	2.9

*SV40-UTR(3.9)-LUC and SV40-UTR(4.1)-LUC were cotransfected with the reference β -galactosidase plasmid pON260 into COS-1 cells as described in *Materials and Methods*. In five separate experiments, LUC activity was measured at 24 and 48 hr posttransfection and normalized to β -galactosidase activity. The numbers represent the ratio between LUC activity measured in cells transfected with the SV40-UTR(3.9)-LUC construct relative to the LUC activity measured in cells transfected with the SV40-UTR(4.1)-LUC construct. nd, not determined.

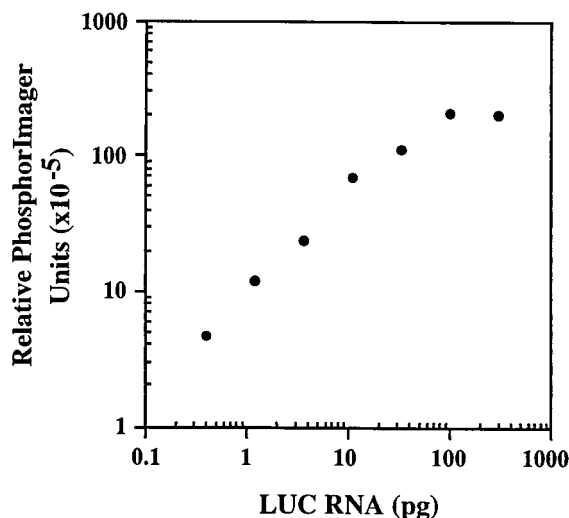


Fig. 4. Standard curve for quantitation of LUC RNA. Samples representing 3-fold serial dilutions of *in vitro*-transcribed LUC RNA (300–0.4 pg) were mixed with 1 μ g of total RNA from nontransfected COS-1 cells. RNA was reverse-transcribed by using primer P3, and PCR was performed by using primers P1 and P2 in the presence of [³²P]dCTP as described in *Materials and Methods*. The PCR products were electrophoresed on an agarose gel and transferred to Nytran. The relative intensities of the bands were measured by using the PhosphorImager. The background signal obtained by using 1 μ g of total RNA from nontransfected COS-1 cells was subtracted from the signal of each sample.

regulated as a result of the switch to the 3.9-kb start site; the use of this start site is governed, in part, by mammary gland-restricted transcription factors (11, 12). In step two, the 3.9-kb mRNA with its truncated 5'-UTR was predicted to be translated more efficiently.

The recognition that the β 4GalT-I gene produces two mRNAs with such dissimilar 5'-UTRs provided the rationale for proposing the second step of the model. We noted the fact that UTR(4.1) is long (only \approx 25% of mRNAs surveyed have 5'-UTRs >150 nt), whereas UTR(3.9) is short (only \approx 5% of mRNAs surveyed have 5'-UTRs <50 nt; ref. 23). Moreover, this sequence is very GC-rich. Thus, UTR(4.1) is predicted to form extensive secondary structure; a feature shown to impair translation. Truncation to a 28-nt 5'-UTR would remove this feature.

In this study, we show that the LUC reporter mRNA containing the short, less-structured UTR(3.9) is translated more efficiently both *in vitro* (\approx 14-fold) and *in vivo* (3- to 5-fold). Consequently, these results confirm and support the translational component of our model and also provide a biological rationale for the switch to the production of the β 4GalT-I transcript with the truncated 5'-UTR. The magnitude of the increase in translational efficiency of the UTR(3.9)-LUC mRNA *in vivo* (3- to 5-fold) is interesting in

Table 2. Quantitation of LUC mRNA in transfected COS-1 cells

Construct	Relative PhosphorImager Units ($\times 10^{-5}$)*	
	Exp. 4	Exp. 5
SV40-UTR(3.9)-LUC	18.4	20.1
SV40-UTR(4.1)-LUC	18.5	21.7

*Total RNA was isolated from transfected COS-1 cells. LUC mRNA was quantitated using RT-PCR as described in *Materials and Methods*. The products were electrophoresed on an agarose gel and transferred to Nytran. The relative intensities of the bands were measured using a PhosphorImager. Each sample was assayed in duplicate, and the averages are reported.

the context of the 50-fold increase in β 4GalT-I activity noted in the lactating mammary gland. A \approx 10-fold increase can be accounted for at the level of transcription, whereas a 3- to 5-fold increase can be accounted for at the level of translation. Our observation that the change in transcription rate is greater in magnitude than the change in translation rate is consistent with observations for other genes under both transcriptional and translational control (24). Thus, as discussed by Mathews *et al.* (24), regulation at the level of translation provides a means for fine control in adjusting cellular levels of a gene product.

Is the Short 5'-UTR Unique to the Mammalian β 4GalT-I Transcript? As noted in the Introduction, all of the mammalian β 4GalT-I genes characterized to date (mouse, human, and cow), express both a 4.1- and a 3.9-kb mRNA. If, as we have proposed (11), the 3.9-kb start site was introduced into the nonmammalian vertebrate β 4GalT-I gene during the evolution of mammals as a direct consequence of the recruitment of this enzyme for lactose biosynthesis, two predictions follow. First, the β 4GalT-I gene from a nonmammalian vertebrate should exhibit a single transcriptional start site to accommodate the role of β 4GalT-I in glycoconjugate biosynthesis. Second, the transcript produced should possess a long, highly structured 5'-UTR similar to that of the mammalian 4.1-kb mRNA.

To test these predictions we have cloned and characterized the β 4GalT gene from the chicken (25). The unanticipated result was the identification of two functional, nonallelic genes (designated CK β 4GalT-I and CK β 4GalT-II); each encodes an α -lactalbumin-responsive β 1,4-galactosyltransferase activity. By using chromosomal mapping results and taking advantage of the comparative gene maps between different species, which reveal regions of evolutionary conserved synteny, we were able to establish that it was the CK β 4GalT-I gene lineage that was recruited from the vertebrate gene pool for the biosynthesis of lactose during the evolution of mammals (ref. 25, see also ref. 26).

As predicted, the CK β 4GalT-I transcript contains both a long (\approx 250 nt), highly structured 5'-UTR and only one possible AUG initiation codon. CK β 4GalT-II also shares the identical structural features. As discussed below, long, highly structured 5'-UTRs may prove to be characteristic of all Golgi-resident glycosyltransferases. This further emphasizes the unique functional role played by the short UTR(3.9) in the tissue-specific biosynthesis of lactose.

Why Is the 5'-UTR of the 4.1-kb β 4GalT-I Long and Highly Structured? The results from this study offer a biological rationale as to why a short, unstructured 5'-UTR is found in the transcript specifically up-regulated in the lactating mammary gland. However, the results do not explain why the 5'-UTR of the 4.1-kb β 4GalT-I mRNA is long and highly structured. Clues come from the observation that other mRNA species known to have long, structured 5'-UTRs encode proteins involved in the regulation of cell growth, differentiation, and development (e.g., oncogenes, transcription factors, growth factors, and their receptors; reviewed in ref. 27). These mRNAs constitute a subset whose translation is preferentially stimulated after growth-factor stimulation of cells in a quiescent (G_0) state. The protein products of these mRNAs are required for progression through the G_1 phase of the cell cycle. Studies to elucidate the signaling pathways involved have shown that control is exerted at the level of translation initiation and depends on the state of phosphorylation of a protein designated 4E-BP1 (28), or PHAS-1 (29).

4E-BP1. Eukaryotic initiation factor 4E (eIF4E; the mRNA cap-binding protein) together with eIF4G (a scaffold protein) and eIF4A (an RNA helicase) form the eIF4F translation initiation complex. As eIF4E is present in the cell in limiting amounts, its availability is thought to control the rate of translation initiation (reviewed in ref. 30). In quiescent cells, eIF4E is bound by 4E-BP1, preventing its interaction with

eIF4G to form eIF4F. However, on growth factor stimulation and reentry into G₁, 4E-BP1 is phosphorylated, causing it to dissociate from eIF4E, thus allowing eIF4E to bind eIF4G (28, 29). Efficient translation of mRNAs containing extensive secondary structure in their 5'-UTRs appears to be sensitive to cellular levels of eIF4E (20); as a result, translation of these mRNAs is quite susceptible to the phosphorylation status of 4E-BP1. Consequently, this subset is preferentially translated after growth factor stimulation (reviewed in refs. 31 and 32).

In retrospect, the data obtained from the *in vitro* translation experiment indirectly supports the idea that translation of the UTR(4.1)-LUC mRNA is more sensitive to eIF4E levels. Initially, we were puzzled by the observation that *in vitro*, the translational efficiency of the UTR(3.9)-LUC mRNA compared with the UTR(4.1)-LUC mRNA was 14-fold higher, whereas *in vivo*, it was 3- to 5-fold higher. However, Rau *et al.* (33) have recently shown that 4E-BP1 is present in reticulocyte lysates and that the predominant form appears to be nonphosphorylated. If 4E-BP1 is unable to be rephosphorylated in these extracts, eIF4F complex formation would be impaired. As a consequence, translation of mRNAs containing long, highly structured 5'-UTRs would be expected to be less efficient *in vitro* than *in vivo*.

The Golgi and Mitosis. Using controls at the level of translation (as opposed to the level of transcription) offers a cell the ability to respond quickly to extracellular signals, thereby avoiding the time lag associated with mRNA synthesis, processing, and transport (24). The signaling pathway involving 4E-BP1 allows for the preferential translation of a subset of mRNAs during progression through the G₁ phase. Could translation of the 4.1-kb mRNA (and other Golgi-resident glycosyltransferase mRNAs) be regulated in a similar manner? During mitosis, the Golgi vesiculates, distributing half of its contents to each daughter cell. If reconstitution of this organelle to its premitotic size occurs as cells progress through G₁, then translation of the mRNAs encoding the glycosyltransferases (and other Golgi-resident proteins) may also be regulated by the 4E-BP1 signaling pathway. Interestingly, an initial inspection of the 5'-UTRs of those glycosyltransferase mRNAs for which a full-length sequence has been published (≈ 10 to date) reveals that each has a long (≈ 200 –600 nt) 5'-UTR that is predicted to be highly structured (unpublished data). This observation suggests that this signaling pathway may, in fact, be involved in the translational control of many, if not all, Golgi-resident glycosyltransferases.

We thank Dr. Bhanu Rajput for discussions and Drs. Curt Civin and David Sabatini (Johns Hopkins University) for the COS-1 and 293 cells, respectively. We also thank the anonymous reviewer for insightful suggestions. This work was supported in part by a National Institutes of Health Grant CA45799 (to J.H.S.). Postdoctoral support to M.C. was provided by a Human Frontiers Science Program Grant RG-414/94 M (to J.H.S.).

1. Beyer, T. A. & Hill, R. L. (1982) in *The Glycoconjugates*, ed. Horowitz, M. (Academic, New York), pp. 25–45.
2. Brodbeck, U., Denton, W. L., Tanahashi, N. & Ebner, K. E. (1967) *J. Biol. Chem.* **242**, 1391–1397.

3. Brew, K., Vanaman, T. C. & Hill, R. L. (1968) *Proc. Natl. Acad. Sci. USA* **59**, 491–497.
4. Hill, R. L., Brew, K., Vanaman, T. C., Trayer, I. P. & Mattock, P. (1968) *Brookhaven Symp. Biol.* **21**, 139–154.
5. Palmiter, R. D. (1969) *Biochem. J.* **113**, 409–417.
6. Turkington, R. W., Brew, K., Vanaman, T. C. & Hill, R. L. (1968) *J. Biol. Chem.* **243**, 3382–3387.
7. Shaper, N. L., Hollis, G. F., Douglas, J. G., Kirsch, I. R. & Shaper, J. H. (1988) *J. Biol. Chem.* **263**, 10420–10428.
8. Russo, R. N., Shaper, N. L. & Shaper, J. H. (1990) *J. Biol. Chem.* **265**, 3324–3331.
9. Masri, K. A., Appert, H. E. & Fukuda, M. N. (1988) *Biochem. Biophys. Res. Commun.* **157**, 657–663.
10. Mengle-Gaw, L., McCoy-Haman, M. F. & Tiemeier, D. C. (1991) *Biochem. Biophys. Res. Commun.* **176**, 1269–1276.
11. Harduin-Lepers, A., Shaper, J. H. & Shaper, N. L. (1993) *J. Biol. Chem.* **268**, 14348–14359.
12. Rajput, B., Shaper, N. L. & Shaper, J. H. (1996) *J. Biol. Chem.* **271**, 5131–5142.
13. Zucker, M. & Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133–148.
14. Pelletier, J. & Sonenberg, N. (1985) *Cell* **40**, 515–526.
15. Kozak, M. (1991) *J. Biol. Chem.* **266**, 19867–19870.
16. Kang, C. & Wu, C. W. (1987) *Nucleic Acids Res.* **15**, 2279–2294.
17. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
18. Suzow, J. & Friedman, A. D. (1993) *Mol. Cell. Biol.* **13**, 2141–2151.
19. Shaper, N. L., Shaper, J. H., Meuth, J. L., Fox, J. L., Chang, H., Kirsch, I. R. & Hollis, G. F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1573–1577.
20. Koromilas, A. E., Lazaris-Karatzas, A. & Sonenberg, N. (1992) *EMBO J.* **11**, 4153–4158.
21. Mathis, D. J. & Chambon, P. (1981) *Nature (London)* **290**, 310–315.
22. Nam, S. C. & Kang, C. W. (1988) *J. Biol. Chem.* **263**, 18123–18127.
23. Pesole, G., Fioramarino, G. & Saccone, C. (1994) *Gene* **140**, 219–225.
24. Mathews, M. B., Sonenberg, N. & Hershey, J. W. B. (1996) in *Translational Control*, eds. Hershey, J. W. B., Mathews, M. B. & Sonenberg, N. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 1–29.
25. Shaper, N. L., Meurer, J. A., Joziassie, D. H., Chou, T.-D. D., Smith, E. J., Schnaar, R. L. & Shaper, J. H. (1997) *J. Biol. Chem.* **272**, 31389–31399.
26. Lo, N.-W., Shaper, J. H., Pevsner, J. & Shaper, N. L. (1998) *Glycobiology* **8**, 517–526.
27. Kozak, M. (1991) *J. Cell. Biol.* **115**, 887–903.
28. Pause, A., Belsham, G. J., Gingras, A. C., Donze, O., Lin, T. A., Lawrence, J. J. & Sonenberg, N. (1994) *Nature (London)* **371**, 762–767.
29. Lin, T. A., Kong, X., Haystead, T. A., Pause, A., Belsham, G., Sonenberg, N. & Lawrence, J. C., Jr. (1994) *Science* **266**, 653–656.
30. Rhoads, R. E. (1993) *J. Biol. Chem.* **268**, 3017–3020.
31. Sonenberg, N. (1996) in *Translational Control*, eds. Hershey, J. W. B., Mathews, M. B. & Sonenberg, N. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 245–269.
32. Lawrence, J. C., Jr., & Abraham, R. T. (1997) *Trends Biochem. Sci.* **22**, 345–349.
33. Rau, M., Ohlmann, T., Morley, S. J. & Pain, V. M. (1996) *J. Biol. Chem.* **271**, 8983–8990.