

# Transcriptional Analysis of the Mouse $\beta$ -Casein Gene

Heather S. Goodman and Jeffrey M. Rosen

Department of Cell Biology  
Baylor College of Medicine  
Houston, Texas 77030

**These studies were designed to further elucidate the relative contributions of transcriptional and post-transcriptional mechanisms involved in  $\beta$ -casein gene regulation in the mammary epithelial cell line designated COMMA-D and a clonal subline designated HC-11. Primary transcripts were mapped under various hormonal and substratum conditions using the technique of nuclear run-on transcription and single stranded sense and antisense probes spanning the  $\beta$ -casein gene. In the presence of insulin alone very little sense transcription is detectable, but antisense transcription is observed, which originates at least 150 basepairs upstream of the normal start site of transcription and is present regardless of hormonal, cell substratum, cell type, or gene activity. Antisense transcription is also detectable in the 3' end of the gene. Insulin, glucocorticoids, and PRL are all necessary for a maximal increase in transcription. A 2- to 4-fold increase in transcriptional activity is observed in the presence of insulin and PRL compared to insulin alone, and this is accompanied by a 125-fold increase in the level of  $\beta$ -casein mRNA. All three hormones act synergistically to induce a 10-fold increase in transcriptional activity, but the transcriptional increase across the gene is not equimolar. The 5' half of the gene is transcribed at a level that is 2- to 10-fold lower than that of the 3' half of the gene. These studies reveal a significant transcriptional component to  $\beta$ -casein gene regulation which was not heretofore detected using double stranded cDNA probes representative of only the 3' half of the gene. (Molecular Endocrinology 4: 1661-1670, 1990)**

## INTRODUCTION

Casein gene expression in explant and cell cultures has been shown to be regulated by a variety of signals, including peptide and steroid hormones and interactions between mammary cells and their substratum (1-3). This regulation has been reported to occur at both the transcriptional and posttranscriptional levels (4, 5). A

synergistic effect of glucocorticoids (F), insulin (I), and PRL (M) has been observed in both explant and cell cultures on  $\beta$ -casein mRNA accumulation (1, 5, 6).

Using a mouse mammary epithelial cell line designated COMMA-D, induction of  $\beta$ -casein mRNA is obtained when cells are grown on floating collagen type I gels in the presence of IFM (7). Run-on transcription assays performed previously in these cells using a double stranded  $\beta$ -casein cDNA probe representative of the 3' end of the gene, indicated that there was a 2- to 3-fold change in transcription which was accompanied by a 60- to 70-fold change in mRNA accumulation (5).

Based on these observations experiments were designed to map primary transcripts in an effort to determine whether they differed under various hormonal conditions and, if so, did this difference correlate with the accumulation of mature  $\beta$ -casein mRNA. To initiate these experiments the mouse  $\beta$ -casein gene was first cloned and characterized, and various repeat-free subclones were generated to use in run-on transcription experiments. Run-on transcription assays essentially measure the density of RNA polymerases present in chromatin at a specific time, and thus, this assay can be used to assess whether there is equimolar transcription in a gene under various conditions. This technique is capable of mapping transcripts that are unstable and difficult to detect in steady state RNA. The use of single stranded probes with this assay has provided information not only about antisense transcription occurring within genes, but also has revealed the mechanism of a block to transcriptional elongation in several genes (8-10).

These studies have revealed several aspects of  $\beta$ -casein gene expression. Strong antisense transcription is detected within the 5' flanking region of the gene very close to the promoter, as well as some weaker antisense transcription at the 3' end of the gene. Additionally, it has been found that there is nonequimolar transcription when the gene is fully induced. The 5' half of the gene is transcribed at a level 2- to 10-fold lower than that of the 3' half. It appears that this difference is due to neither RNA processing nor the presence of an internal promoter. [ $^3$ H]Uridine pulse labeling studies in intact cells revealed findings similar to those observed with the nuclear run-on transcription assay.

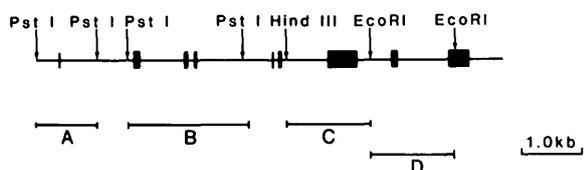
## RESULTS

### Isolation of Genomic Clones

To map primary transcripts using the nuclear run-on assay homologous genomic clones were needed. Since available cell culture lines displaying hormonally regulated  $\beta$ -casein gene expression are mouse derived, a mouse genomic library was screened using previously characterized rat  $\beta$ -casein genomic probes, and several  $\lambda$ -phage clones were isolated. One such clone could be digested with *EcoRI* to yield three fragments which were 1.55, 6.6, and 1.35 kilobases (kb) in length. These fragments were subcloned into plasmid vectors and restriction mapped, and portions of each were sequenced. The 6.6-kb fragment contained the majority of the gene, the 1.55-kb fragment consisted of 5' flanking DNA, and the 1.35-kb contained the 3' end of the gene. Since the 6.6-kb fragment encompassed most of the gene it was further subcloned and sequenced in more detail. All of the exon/intron boundaries were located, and approximately 400 basepairs (bp) of 5' flanking DNA were also sequenced. The sequence derived from this clone as well as the 5' flanking region is in agreement with that reported by Yoshimura and Oka (11). Various repeat-free subclones were generated which spanned the locus of the gene. The presence of repeats was tested by labeling each subcloned DNA fragment and using it to probe genomic blots. Each final subclone selected gave a single band on Southern blots of the predicted size. Four repeat-free clones were selected, which are diagrammed in Fig. 1, and then used in nuclear run-on analysis.

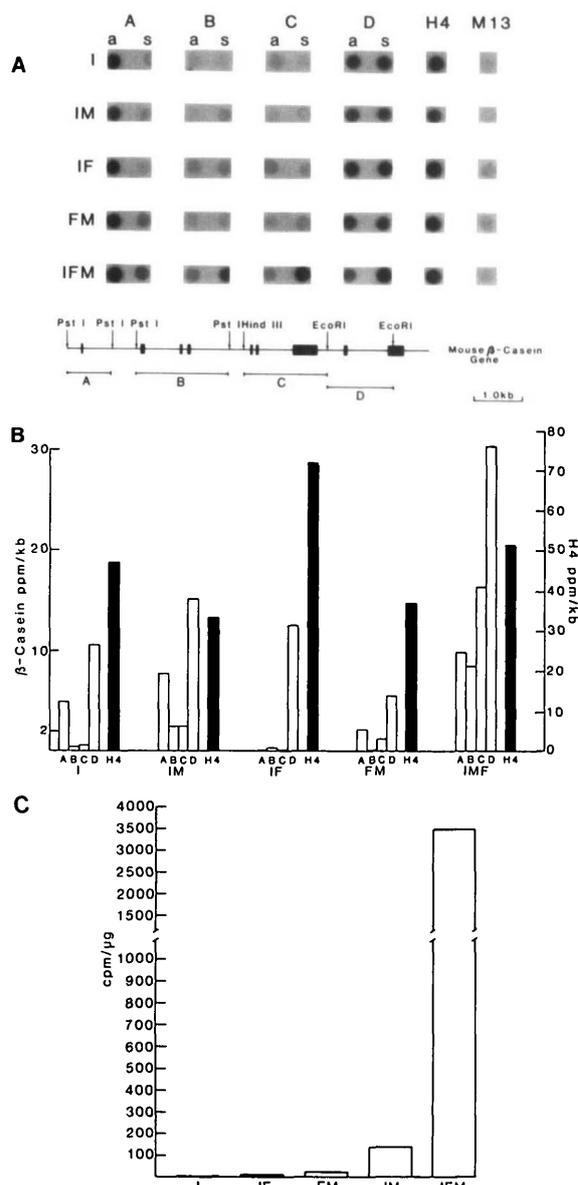
### Nuclear Run-On Analysis Using Strand-Specific Clones

The four clones were subcloned into M13 vectors, and strand-specific probes were generated. A series of run-on assays was performed using nuclei isolated from COMMA-D cells grown on floating type I collagen gels under a variety of hormonal conditions, as shown in Fig. 2. The results from one such experiment are depicted in the autoradiogram shown in Fig. 2A. The levels of sense transcription are summarized in Fig. 2B, normalized for counts per min input, clone size, and hybridization efficiency. Antisense transcription is design-



**Fig. 1.** Map of the 6.6-kb Fragment Used in Run-On Transcription Experiments

This diagram illustrates various restriction sites and the exon and intron positions within the 6.6-kb fragment. Below the diagram the four repeat-free subclones, A-D, are designated which span the gene.



**Fig. 2.** Run-On Transcription in COMMA-D Cells Grown Under Various Hormonal Conditions

COMMA-D cells were grown under various hormonal conditions on floating collagen gels for 48 h. After this time, either nuclei were isolated and used in run-on transcription assays or total RNA was isolated and used to quantitate steady state  $\beta$ -casein RNA levels. A is an autoradiogram of nuclear run-on transcription results. Clones A-D were contained in single stranded M13 vectors. Antisense transcription is designated a, and sense transcription s. H4 is a histone H4 control which is in a double stranded plasmid vector. M13 is M13 vector alone, used to determine the nonspecific background. B is the quantitation of sense transcription shown in A. The values have been normalized with respect to background signal, clone size, number of input counts, and hybridization efficiency. C is a quantitation of  $\beta$ -casein steady state RNA which was determined by slot blot analysis followed by scintillation counting.

nated a, and sense transcription as s in Fig. 2A. For each condition a mouse histone H4 filter is also included as a control for a gene that is not hormonally responsive in these cells. The M13 vector alone filter is also included to assess nonspecific background. For each condition a parallel plate of cells was used to isolate total RNA in order to determine the steady state level of  $\beta$ -casein RNA. These results are depicted in Fig. 2C.

This set of experiments illustrates several unexpected aspects of  $\beta$ -casein gene expression. First antisense transcription observed originating from the A clone is present at the same intensity regardless of the hormonal condition. A weaker antisense transcriptional signal as well as some sense transcription are also detected within the D clone under all of the hormonal conditions monitored. However, in the presence of IMF, sense exceeds that of the antisense transcription from the D region.

Interestingly, in the presence of IMF there are significant increases in transcription across the gene, but these increases are not equivalent. The 5' half of the gene represented by clones A and B gives a signal that is half as large as that observed for the 3' half of the gene, as represented by clones C and D (Fig. 2B). Therefore, there appears to be nonequimolar transcription across the gene even when the gene is fully induced in these cells.

Furthermore, the presence of all three hormones is necessary for a significant increase in transcription across the gene. A small increase (1- to 3-fold) is apparent in the presence of IM (Fig. 2B). This transcriptional increase is at most 3-fold if compared using the B or C probe and considerably less when comparing the signals obtained with the A and D probes. This increase in the relative rate of transcription can be contrasted with the changes in the steady state level of  $\beta$ -casein mRNA determined in parallel under these hormonal conditions (Fig. 2C). In the presence of IM there is minimally a 125-fold induction of  $\beta$ -casein mRNA compared to I alone (assuming the basal level to be 1). Thus, the increase in transcription from I to IM is too small to account for the increase in steady state  $\beta$ -casein RNA levels. M might, however, have induced an early and transient increase in transcription, which disappears by 48 h, at which time the levels of steady state mRNA and transcription were assayed. This possibility was tested by performing a kinetic experiment in which run-on transcription assays were conducted 1, 4, 8, and 24 h after M addition. No early effect on transcription was found (data not shown). Thus, based upon these results the role of M in the presence of I appears to be predominantly posttranscriptional.

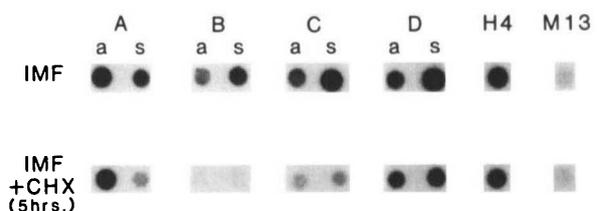
F in combination with I do not lead to an increase in transcription, as illustrated by the lack of a significant change in the relative rate of transcription from I to IF (Fig. 2, A and B). M in combination with I stimulates a small increase in transcription, and upon the addition of F this level is further induced. This series of experiments clearly demonstrates that I, F, and M act synergistically in activating  $\beta$ -casein gene transcription.

### Cycloheximide Treatment Inhibits $\beta$ -Casein Gene Transcription

Previous experiments were designed to determine whether hormones were acting directly or indirectly to induce  $\beta$ -casein gene expression by examining the effects of cycloheximide treatment on  $\beta$ -casein mRNA accumulation in COMMA-D cells. Addition of cycloheximide for 5 h to COMMA-D cells after 48 h in the presence of IMF had little effect on  $\beta$ -casein mRNA levels. However, if cells were incubated with IM for either 24 or 48 h or IMF for 24 h cycloheximide treatment led to a dramatic reduction in the level of  $\beta$ -casein mRNA (27). Additionally, either cycloheximide or puromycin treatment has been shown to block the hormonally induced increase in  $\beta$ -casein transcripts in mouse mammary explants (35). Therefore, it was of interest to know if this decrease was in part due to a transcriptional effect. To determine whether there were any labile factors that were necessary for transcription of the gene, cells were exposed to cycloheximide for 5 h after 48 h of pretreatment with IMF. Labile transcription factors or posttranscriptional modifications of these factors whose presence depends on ongoing protein synthesis should be sensitive to cycloheximide treatment. As shown in Fig. 3 treatment with cycloheximide significantly reduces (by >95% in the B and C clones) the level of sense transcription across the gene. The pattern of transcription after cycloheximide treatment resembles that found in the presence of I alone. Interestingly, the level of antisense transcription detected with the A and D probes is unaffected by this inhibitor. Therefore, some transcription factor (or factors) is either induced by I, M, or F, or a posttranscriptional modification of this factor(s) is dependent upon ongoing protein synthesis.

### The 5' Antisense Transcription Is Due to RNA Polymerase II Activity, Is Present in Nonexpressing Cells, and Is Confined to 5' Flanking DNA

Since the 5' antisense transcription signal was quite high and was occurring in a conserved portion of the gene near the promoter, several additional experiments

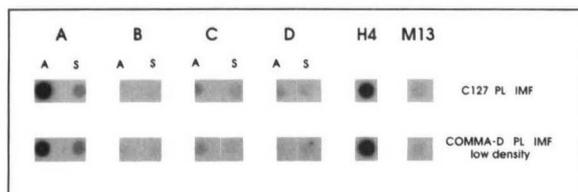


**Fig. 3.** Effect of Cycloheximide on  $\beta$ -Casein Transcription

COMMA-D cells were grown in the presence of IMF for 48 h and then either treated with cycloheximide (CHX) for 5 h or left untreated. Nuclei were isolated, and run-on transcription was performed as described. The corrected parts per min/kb for sense transcription were as follows: IMF: A = 14.6; B = 11.4; C = 65.5; D = 50.7. IMF + CHX: A = 4.8; B = 0; C = 3.6; D = 11.8.

were performed to assess whether antisense transcription played any role in the regulation of  $\beta$ -casein gene sense transcription. In cases where antisense transcripts have been shown to be involved in gene regulation, their role has usually been inhibitory (14–16). However, it is difficult to envision a negative role in this case because the level of antisense transcription is the same regardless of the activity of the gene in the COMMA-D cells (Fig. 2B). A high level of antisense transcription was also observed in cell types that do not express  $\beta$ -casein (Fig. 4). C127 cells are a mammary cell line that does not express  $\beta$ -casein, but exhibits a comparable level of antisense transcription. COMMA-D cells also produce barely detectable levels of  $\beta$ -casein mRNA when grown on plastic at low density regardless of the presence of lactogenic hormones, yet a level of 5' antisense transcription similar to that detected previously is observed (Fig. 4). Antisense transcription was also detected in NIH 3T3 cells, a fibroblast cell line (data not shown). Thus, the 5' antisense transcription is observed using nuclear run-on assays regardless of the hormonal or substratum condition and regardless of cell type. Interestingly, in COMMA-D cells grown at low density on plastic no signal was detected, either sense or antisense, from the D clone. Thus, it is possible that 3' antisense transcription is somehow related to cell-substratum interactions.

It is conceivable that the 5' antisense transcription might inhibit sense transcription, but this was not detectable in the COMMA-D cells. Only about 10–20% of the COMMA-D cells express  $\beta$ -casein. Thus, if 5' antisense transcription was a mechanism to silence the gene, then even in the presence of IMF 80–90% of the cells might still be expressing the antisense transcription. Accordingly, there would not be a significant change in the intensity of this signal under the different hormonal conditions studied in COMMA-D cells. To examine this possibility a clonal cell line, designated HC-11, derived from the COMMA-D cells was employed (17). Under the appropriate hormonal conditions approximately 90% of these cells express  $\beta$ -casein (Doppler, W., and R. K. Ball, personal communication). A distinguishing feature of these cells is their loss of the



**Fig. 4.** Transcriptional Analysis of  $\beta$ -Casein Gene Expression in a Nonexpressing Cell Line vs. That in COMMA-D Cells Grown on Plastic (PL)

Cells were grown on plastic in the presence of IMF for 48 h. The COMMA-D cells were maintained at a low density, which is not conducive to  $\beta$ -casein gene expression. Nuclei were isolated and used in run-on transcription assays as described. Forty-three  $\times 10^6$  cpm were used in the C127 cell hybridization and 54  $\times 10^6$  cpm for the COMMA-D cells.

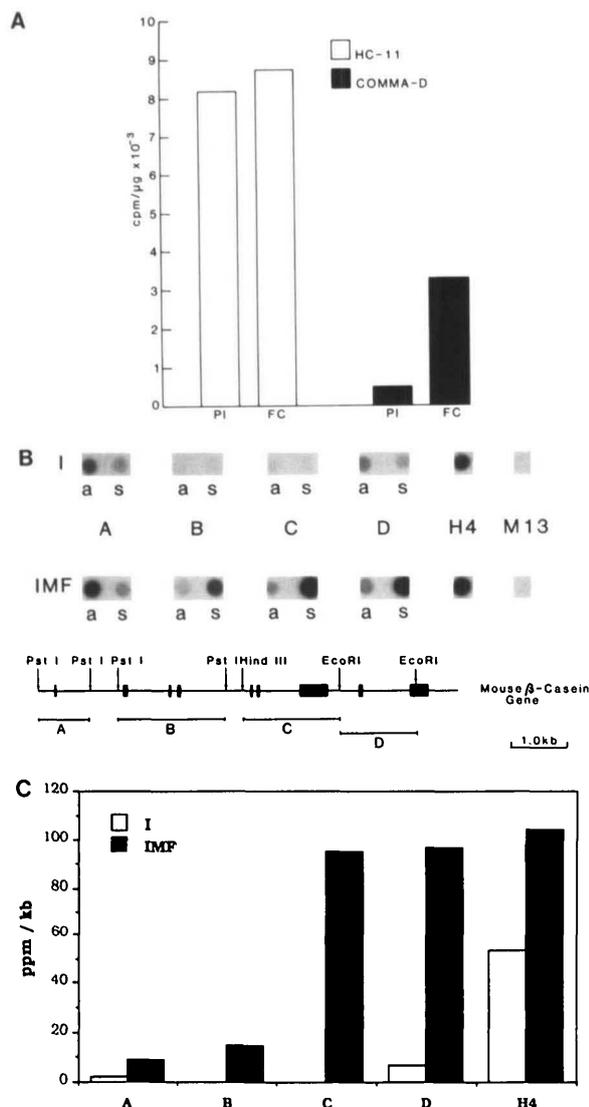
cell substratum requirement for  $\beta$ -casein gene expression, as shown in Fig. 5A. This difference in cell substratum dependence compared to that of the parental COMMA-D cells is apparent, in that there is approximately a 5- to 7-fold increase in the steady state level of  $\beta$ -casein mRNA in the COMMA-D cells when they are grown on floating collagen compared to that in cells grown at high density on plastic in the presence of IMF. In contrast, the HC-11 cells display no cell substratum dependence of  $\beta$ -casein gene expression, and approximately twice the level of  $\beta$ -casein mRNA is detected compared to that in the COMMA-D cells. The transcriptional pattern in HC-11 cells is very similar to that observed in the COMMA-D cells in the presence of either I alone or IMF (Fig. 5B). Thus, it is unlikely that nonexpressing cells could be responsible for the 5' antisense transcriptional signal observed. The level of sense transcription in HC-11 cells is quantitated in Fig. 5C, illustrating once again the nonequimolar transcription across the gene. Interestingly, the gradient across the gene is even greater in these cells than in COMMA-D cells.

It was of interest to know whether the 5' antisense was originating within the gene and, therefore, generating an internal antisense transcript capable of potentially forming double stranded RNA, or if it originated in the 5' flanking region, and the promoter was bidirectional. Therefore, the A clone was subdivided into four *Rsa*I fragments, some of which contained either 5' flanking or intragenic sequences. These subclones were used in run-on transcription experiments to map the precise location of the 5' antisense transcription. As illustrated in Fig. 6A, antisense transcription is confined to the 5' flanking DNA and originates at least 150 bp up-stream of the normal start site of transcription. Furthermore, this positioning does not appear to be hormonally regulated, since it did not change depending on whether the cells were grown in I or IMF.

To ascertain if the antisense signal was due to RNA polymerase II activity a run-on transcription experiment was performed in the presence of 1  $\mu$ g/ml  $\alpha$ -amanitin, a concentration that will specifically inhibit RNA polymerase II activity, but not RNA polymerase I or III activity. Both the 5' and 3' antisense signals in addition to the sense transcription are  $\alpha$ -amanitin sensitive (Fig. 6B). A 28S ribosomal RNA clone is included as a positive control in this experiment to demonstrate the presence of RNA polymerase I transcription in the presence of this concentration of  $\alpha$ -amanitin. Therefore, the antisense transcription observed is due to RNA polymerase II activity.

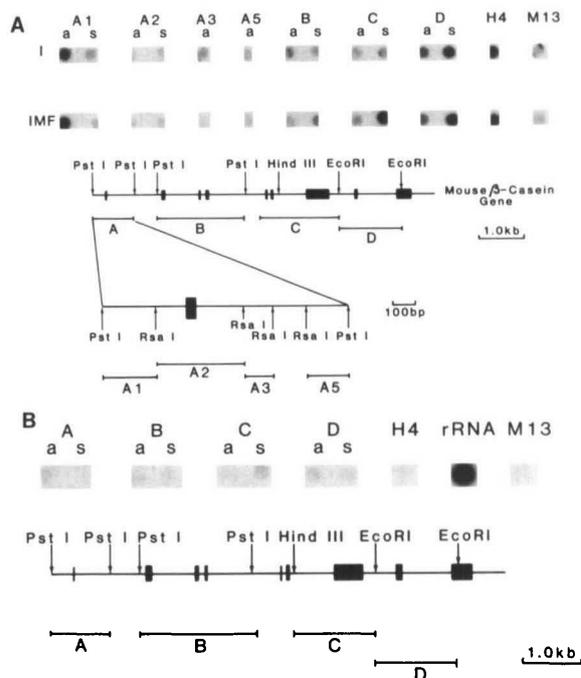
#### Nonequimolar Transcription Is not Due to RNA Processing

The nonequimolar transcription observed in the presence of IMF could conceivably have been the result of an additional hormonally responsive promoter located within the region encompassed by the C clone or located somewhere between clones B and C. To test



**Fig. 5. Beta-Casein Gene Expression in HC-11 Cells**  
 A illustrates the comparison of steady state levels of  $\beta$ -casein RNA in COMMA-D and HC-11 cells. Cells were grown in the presence of IMF on either plastic (PI) or floating collagen (FC). Both HC-11 and COMMA-D cells were grown at high density (confluence) for several days before hormone addition. After 48 h of hormone treatment with IMF total RNA was isolated and quantitated by slot blot analysis. B depicts the transcriptional analysis of  $\beta$ -casein gene expression in the HC-11 cells. HC-11 cells were grown on plastic at confluence for 4 days in the presence of I and stripped horse serum and then in the presence of IMF for 48 h, as described in *Materials and Methods*. Nuclei were isolated and used in run-on transcription assays as previously described. C shows the quantitation of the sense transcription from the experiment in B. Values are normalized with respect to vector background signal, clone size, number of input counts, and hybridization efficiency.

this possibility a series of Northern analyses was performed using the B, C, and D clones as probes (data not shown). If there was an internal promoter, another RNA transcript besides that representing the mature  $\beta$ -casein mRNA should be observed, assuming that this transcript was stable. However, the only transcript de-



**Fig. 6. Antisense Transcription Is Confined to the 5' Flanking DNA and Is Due to RNA Polymerase II Activity**

A illustrates that the 5' antisense transcription is confined to the 5' flanking DNA. COMMA-D cells were grown with either I or IMF for 48 h on floating collagen gels. Nuclei were isolated, and run-on transcription assays were performed. The A clone was divided into four *Rsa*I fragments, which are designated in the *bottom diagram* as A1, A2, A3, and A5. B shows that  $\alpha$ -amanitin inhibits both sense and antisense transcription. COMMA-D cells were grown for 48 h in the presence of IMF on floating collagen gels. Nuclei were isolated, and run-on transcription was performed in the presence of 1  $\mu$ g/ml  $\alpha$ -amanitin. A 28S ribosomal RNA filter served as a positive control (rRNA).

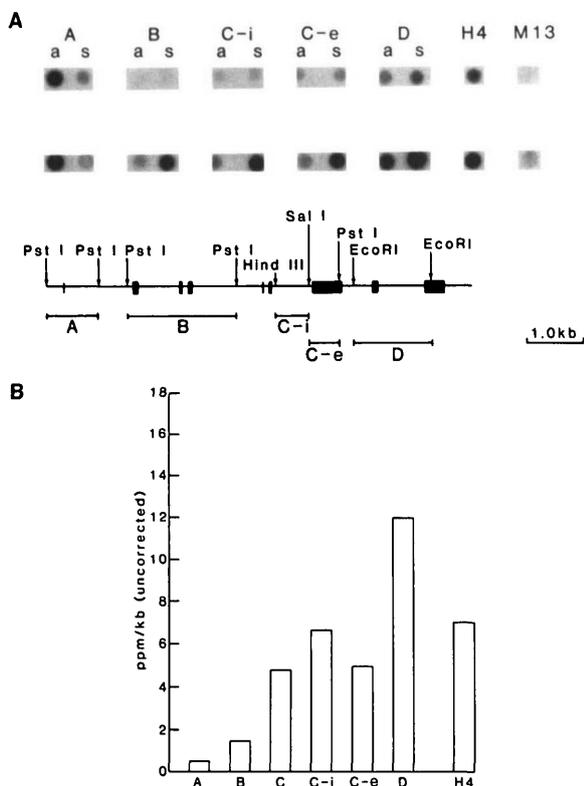
tected was  $\beta$ -casein mRNA of the expected size. Therefore, it seems unlikely that an internal promoter can account for the unequal transcriptional signal across the gene, but this possibility cannot be ruled out at this point.

An alternative explanation for this phenomenon is that RNA splicing and selective intron turnover might occur in the isolated nuclei, resulting in unequal signals across the gene in the presence of lactogenic hormones, because the 5' half of the gene is very intron rich compared to the 3' half. Therefore, if the introns were being spliced out and turned over during the 20-min period of the run-on transcription assay, this could account for the lower signal observed from the 5' half of the gene. To test this possibility intron- and exon-specific probes were generated. If splicing and selective turnover were occurring, then the intron probe should give a markedly reduced signal, whereas the exon probe would be expected to yield an enhanced signal. If selective RNA turnover was not occurring, then the signals should be equivalent when expressed as parts per million/kb. The results from such a nuclear run-on

experiment performed in HC-11 cells are shown in Fig. 7A. The HC-11 cells were employed in this experiment rather than COMMA-D cells, because they displayed even a larger gradient of nonequimolar transcription across the gene (Fig. 5C). Intron and exon probes, designated C-i and C-e, respectively, were generated from the C clone. As shown in Fig. 7A, the signal from the intron probe is quite strong. Furthermore, when the results are quantitated as summarized in Fig. 7B, it is apparent that the signals from the exon and intron probes are equivalent. Therefore, it does not appear that selective intron turnover can account for the difference in transcription across the gene in the presence of IMF.

**Pulse Labeling of Cells Results in a Transcription Pattern Similar to That Obtained with Nuclear Run-On Transcription**

To ensure that results obtained from isolated nuclei in run-on transcription assays were not artifacts of the *in*



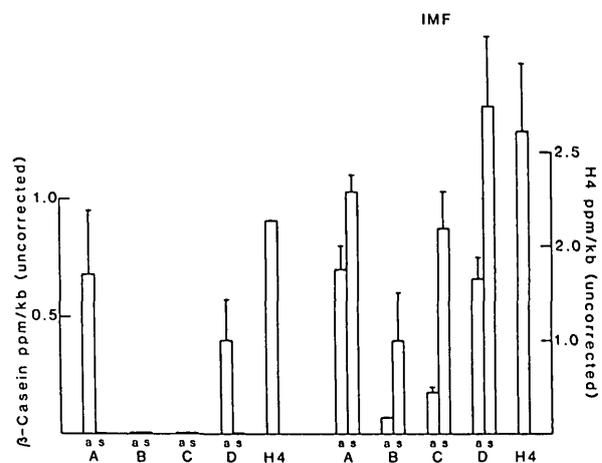
**Fig. 7.** Nonequimolar Transcription Is not Due to RNA Processing

In A, HC-11 cells were grown on plastic in the presence of IMF for 48 h as described. Nuclei were isolated, and run-on transcription was performed. The intron probe designated C-i (870 bp) and the exon probe C-e (505 bp) were both derived from probe C. Their positions are illustrated in the diagram. B is a quantitation of the sense transcription depicted in A. Values are normalized with respect to background signal, clone size, and number of input counts. These results were not, however, corrected for hybridization efficiency. Accordingly, the absolute values are lower than those reported in Fig. 2.

*vitro* system, [<sup>3</sup>H]uridine pulse assays were also performed. These experiments could not be performed in COMMA-D cells due to the necessity of growing the cells on collagen gels for  $\beta$ -casein gene induction. The substratum made it impossible to incorporate sufficient isotope into the cells in a pulse label to accurately measure the levels of  $\beta$ -casein gene transcription. HC-11 cells were, however, well suited for this type of experiment because of their ability to express high levels of  $\beta$ -casein mRNA when grown on plastic (Fig. 5A).

In these experiments, HC-11 cells were given a 30-min pulse of [<sup>3</sup>H]uridine after 48 h of hormonal treatment. Nuclei were isolated, and the RNA analyzed. The results are depicted in Fig. 8. The values are not corrected for hybridization efficiency and are, therefore, much lower than those values obtained in the nuclear run-on transcription experiments. The absolute values in this experiment cannot accurately be compared to those in the nuclear run-on transcription experiments due to the inherent differences in the two experiments, such as the specific activities of the isotopes used and the efficiency of transcription in isolated nuclei vs. intact cells. However, the distribution of transcription across the gene and the ratios between various clones can yield meaningful comparisons.

The results of this experiment are essentially in agreement with those obtained using the run-on transcription. A strong antisense transcription signal is again observed using the A clone in the presence of both I and IMF. Furthermore, an antisense transcription signal was also detected using the D clone. Nonequimolar transcription is still observed across the gene in the presence of IMF, as revealed by the higher signal from the 3' half compared to the 5' half of the gene. In the pulse-labeling experiment, however, the relative signals from



**Fig. 8.** [<sup>3</sup>H]Uridine Labeling in HC-11 Cells

HC-11 cells were grown for 48 h in the presence of either I or IMF on plastic. After 48 h, cells were incubated with [<sup>3</sup>H]uridine for 30 min. Nuclei were isolated, and RNA were extracted and analyzed for  $\beta$ -casein transcription as previously described. Results are normalized with respect to background signal, clone size, and number of input counts. The results are not normalized for hybridization efficiency.

the A and B clones are increased compared to those in the nuclear run-on transcription experiments. This most likely results from the initiation of transcription occurring in intact cells, which does not occur appreciably in isolated nuclei, and illustrates that RNA polymerase transcribes this region of the gene in intact cells. However, it is still unclear why the signal observed in the 5' region of the gene is so low in isolated nuclei.

The results of the pulse-labeling experiments, in general, support the conclusions obtained from the nuclear run-on transcription experiments. Thus, both the presence of antisense transcription and the nonequimolar transcription across the gene observed in isolated nuclei occur in intact cells and are not due to some manifestation of isolated nuclei.

## DISCUSSION

An analysis of mouse  $\beta$ -casein gene transcription has been performed in mammary epithelial cell lines using a series of single stranded genomic clones that span the locus of the gene. Previous experiments designed to analyze the relative rate of  $\beta$ -casein gene transcription relied on the use of either explant or primary culture systems and double stranded cDNA probes. The development of mammary epithelial cell lines that express the endogenous  $\beta$ -casein gene in a hormonally and cell substratum-regulated fashion has facilitated both gene transfer experiments as well as studies of the endogenous gene regulation. The use of these cell lines makes it possible to isolate large quantities of nuclei in order to perform the nuclear run-off assays and, thus, recover sufficient amounts of [ $^{32}$ P]RNA ( $5.0 \times 10^7$  to  $1.0 \times 10^8$ ) to accurately measure transcription rates in both the induced and uninduced states. The use of single stranded probes in these assays has provided further information which has clarified seemingly conflicting results from several laboratories (1, 4, 5, 18, 19, 20) about whether  $\beta$ -casein gene expression is controlled primarily at the transcriptional or posttranscriptional level.

Double stranded cDNA clones used previously in nuclear run-on transcription experiments represent only the last three exons of the gene contained in the C and D clones used in these experiments. Antisense transcription as well as some sense transcription is detected in the D clone, which most likely accounts for the signal obtained from the cDNA in the presence of insulin alone. This previously led to the underestimation of the extent of the transcriptional component of  $\beta$ -casein gene regulation. However, the remainder of the gene exhibits essentially no sense transcription in the presence of insulin alone. Therefore, these experiments indicate that lactogenic hormones exert a profound effect on  $\beta$ -casein gene transcription.

These studies also demonstrate that the synergistic action of all three lactogenic hormones is necessary for a significant increase in transcription of the  $\beta$ -casein

gene. The role of M in  $\beta$ -casein gene regulation is both transcriptional and posttranscriptional. The addition of M to cells in the presence of insulin leads to only a 2- to 4-fold increase in transcription, while there is approximately a 125-fold increase in the steady state level of mRNA (Fig. 2, B and C). Based upon the analysis of casein-chloramphenicol acetyltransferase (CAT) constructs transfected into HC-11 cells, Doppler *et al.* (21) conclude that M action is primarily transcriptional. However, upon addition of M alone these investigators also reported only a 4-fold induction of CAT activity over the basal level obtained in the absence of hormones. This increase is similar to the transcriptional increase detected using nuclear run-on experiments upon the addition of M to cultures containing insulin alone. What was not analyzed in these experiments was the hormonal induction of  $\beta$ -casein mRNA levels in HC-11 cells in parallel experiments. Even so, it would probably be invalid to compare the level of CAT enzyme activity and steady state CAT RNA with that of  $\beta$ -casein RNA. Thus, any posttranscriptional component cannot be analyzed by this means.

Additionally, Doppler *et al.* (21), using the same experimental design, reported a lack of a requirement for insulin in the hormonal induction of  $\beta$ -casein promoter-driven CAT activity in the HC-11 cells. This is in contrast to our results in COMMA-D cells, in which there is a strict insulin requirement for  $\beta$ -casein gene transcription (Fig. 2, A and B). This discrepancy may be explained by the presence of trace amounts of insulin present in the serum used in the casein-CAT transfection experiments in HC-11 cells or the failure of this cell variant to require insulin for  $\beta$ -casein gene expression. Previous studies in rat mammary gland explant cultures have demonstrated that insulin is essential for  $\beta$ -casein gene transcription and does not affect the stabilization of the casein mRNA (18). Recently, Yoshimura and Oka (20) reported that significant expression of casein-CAT fusion genes in transfected primary mouse mammary epithelial cells is dependent upon the presence of insulin.

The primary role of F in  $\beta$ -casein gene expression appears to be indirect (21, 21a). The addition of hydrocortisone in combination with insulin results in extremely low levels of transcription, as does the combination of hydrocortisone with M. Steady state levels of  $\beta$ -casein mRNA are essentially undetectable under these conditions as well. For F to have an effect on transcription or steady state RNA accumulation it is necessary to include both I and M. This is unlike the combination of I and M, which results in a small but significant elevation in transcription, and a larger induction of  $\beta$ -casein mRNA levels. There is some support for the hypothesis that glucocorticoids act at least in part by stimulating the accumulation of some factor(s) that is necessary for  $\beta$ -casein mRNA accumulation (12). Whether this same pathway is used in stimulating  $\beta$ -casein gene transcription is unclear at present.

The large reduction in sense transcription produced by the addition of cycloheximide to the cells after hor-

monal induction is similar to the results obtained by Yoshimura and Oka (13) using [<sup>3</sup>H]uridine pulse studies in explant cultures. This result suggests the presence of at least one, if not more, labile transcription factor or a posttranslational modification activity sensitive to the inhibition of protein synthesis, which is specific for sense transcription. Antisense transcription is not affected under these conditions.

Run-on transcription experiments and pulse labeling in intact cells demonstrated antisense transcription at both the 5' and 3' ends of the gene. Antisense transcription has been found in other genes (15, 22) and may be found to be quite common if experiments are designed appropriately to detect it. When antisense RNA has been shown to have a role in gene regulation it is usually an inhibitory effect (15). The 5' antisense observed in the  $\beta$ -casein gene is invariant with respect to any hormonal or substratum condition tested and is present in cells that do not express  $\beta$ -casein. Both the 5' and 3' antisense transcriptions are due to RNA polymerase II activity. The 5' antisense transcription is confined to the 5' flanking region of the gene, and the promoter, therefore, appears bidirectional. However, Northern analysis so far has failed to detect any stable transcript resulting from either 5' or 3' antisense transcription (Goodman, H. S., and J. M. Rosen, unpublished observation). The 3' antisense transcription must occur within the gene, since the D clone contains the last intron and a portion of the last exon. If this antisense transcription is in any way regulatory, then the mechanism of regulation must be of an inherently different nature than those described previously (14, 15) in which antisense transcripts are readily detectable in steady state RNA. It is unlikely that the antisense transcription detected arises from an unrelated gene that shares homology with the  $\beta$ -casein gene, because all of the probes used in these assays produce a single band upon Southern analysis of genomic DNA, suggesting that they represent a single copy gene. Also, the run-on transcription protocol includes a RNase treatment step such that if the antisense signals were from another gene, the sequences would have to be rigorously basepaired to withstand this part of the procedure.

In the induced state equimolar transcription across genes is usually observed, but this has only been investigated carefully for a limited number of genes (8, 23, 24). One notable exception is the accumulation of RNA polymerase II near the termination site well downstream from the site of polyadenylation, which is thought to account for an increased signal in nuclear run-on experiments using the ADA gene (25). In the studies reported in HC-11 cells a similar distribution of RNA polymerase II molecules is observed in the intact cells using [<sup>3</sup>H]uridine pulse experiments as in the nuclear run-on assays. Additionally, similar results have been obtained in nuclear run-on transcription assays performed in mouse primary epithelial cells cultured on an EHS matrix (Bissel, M., personal communication). One explanation for the nonequimolar distribution of transcription is that pausing of RNA polymerase II is

occurring somewhere within the 3' half of the gene or between the 5' and 3' halves of the gene. This has been shown to be a means of regulation in a variety of genes, but in most cases pausing usually occurs very near the start site of transcription or within the first intron (8, 23, 24, 26-30). To test this idea a series of run-on transcription experiments were performed in the presence of sarkosyl, which if used in the proper concentration will strip loosely associated proteins from chromatin (data not shown). Presumably, this might release a block to RNA polymerase II elongation, and equimolar transcription would then be observed across the gene. This was not observed, however, suggesting that pausing may not account for the nonequimolar transcription observed. Additionally, if pausing were occurring, one would expect at some point in time after hormonal addition to have observed increased transcription of the 5' half of the gene. Various kinetic experiments have failed to detect this transcriptional increase.

A second possibility to explain the nonequimolar transcription is that there is less efficient elongation of RNA polymerase II through the 5' half of the gene. This could be due to a different chromatin conformation in this region. It is possible that some elongation factor is necessary to efficiently transcribe the 5' half of the gene that is not present in these cells, perhaps because of the culture conditions employed. One possibility is that a cell-substratum component is missing in these cells which is necessary for the induction of this putative elongation factor. In *E. coli*, elongation of transcription is a common mechanism of gene regulation, an example being the nusA elongation factor. Finally, the presence of an internal promoter cannot be ruled out, although there is not any direct evidence to support this possibility.

The experiments described here have served to clarify some differences reported previously concerning the mechanisms of casein gene regulation as well as reveal some new information important in casein gene expression. The previous conflict about whether casein regulation was predominantly transcriptional or posttranscriptional was most likely a result of the nature of probes used in nuclear run-on transcription experiments. We have found that it is critical to use single stranded probes that span the gene in this analysis. Antisense transcription may not be a rare occurrence in genes, and the only way to assess its contribution in run-on transcription assays is to use single stranded probes. Additionally, Zilz *et al.* (31) have reported differences in results of experiments designed to assess the effects of thyroid hormone depending whether single or double stranded probes were used in run-on transcription experiments, although this was not apparently due to antisense transcription. The fact that the entire gene may not be transcribed at the same rate also illustrates the importance of using probes that correspond to different areas of the gene.

Although these studies were directed at examining  $\beta$ -casein transcription patterns, probably a more pow-

erful contribution to regulation of casein gene expression is made by the intrinsic casein mRNA stability. The long half-life of casein mRNA appears to contribute to the absolute level of mRNA accumulation and protein synthesized to a greater extent than the absolute rate of transcription. The transcription rate for  $\beta$ -casein even in the fully induced state is not appreciable. For example, it is not as high as that observed for the rate of histone H4 gene transcription, and yet casein mRNA levels are at least an order of magnitude greater than histone H4 levels in the mammary cells. The casein gene produces a very stable transcript in the presence of the proper hormonal and cell substratum conditions, with a half-life of more than 24 h. Studies are in progress to try to identify regions of the mRNA that may confer its intrinsic stability and may possibly play a role in the regulation of casein gene expression.

## MATERIALS AND METHODS

### Phage Library Screening

A C57/B16 mouse genomic library was kindly provided by Dr. John Seidman at Harvard Medical School. This library was constructed as a partial *Hae*III digest with ligated *Eco*RI and cloned into the Charon 4A phage. Library screening was performed using rat genomic and mouse  $\beta$ -casein cDNA probes, as previously described (32, 33), to obtain the mouse  $\beta$ -casein gene. Positives from the initial screen rescreened in secondary and tertiary screens until they were purified. These were then digested with *Eco*RI and subcloned into pGEM Blue vectors (Promega, Madison, WI) for further analysis.

### Clone Characterization

The mouse  $\beta$ -casein gene was characterized by restriction mapping and hybridization to various rat genomic clones (33). Certain subclones were sequenced using the Sequenase system. Selected clones were then subcloned into M13mp8 or mp9 single stranded phage, as previously described (34), and were analyzed for orientation by sequencing.

### M13 Preparation

M13 phage were grown essentially as previously described (34). For large scale DNA preparations, the length of incubation was found to be critical to prevent double stranded DNA contamination. Incubations were carried out for 6 h, after which time the infected cells were removed by centrifugation for 20 min at 8000 rpm at 4 C. Phage were then precipitated by adding 0.1 vol 5 M NaCl and 0.1 vol 50% polyethylene glycol 6000. Precipitation was performed overnight at 4 C. The solution was then pelleted by centrifugation at 8000 rpm for 30 min at 4 C and resuspended in 5% of the original culture volume in TE buffer (10 mM Tris, pH 7.4, and 1 mM EDTA) containing 0.5% Sarkosyl. Phage were precipitated a second time and resuspended in 1% of the original culture volume. Phage DNA was isolated by extracting the particles in phenol, phenol/chloroform, and chloroform.

### Cell Culture

Early passage COMMA D or HC11 cells (passage numbers less than 12; gifts from Dr. Daniel Medina, Baylor College of Medicine, Houston, TX, and Dr. Roland Ball, Friedrich Miescher Institute, Basel, Switzerland, respectively) were

grown essentially as described previously (1, 9). COMMA-D cells of passage 10 or less were grown in Dulbecco's Modified Eagle's Medium (Gibco, Grand Island, NY) supplemented with glutamine (Sigma), 5% fetal calf serum (Hazleton), 5  $\mu$ g/ml I (Sigma), and 50  $\mu$ g/ml gentamycin sulfate (Sigma). HC-11 cells were grown in RPMI medium (Hazleton) with 1 ng/ml epidermal growth factor (Sigma) included as well. Before hormone treatment cells were incubated in medium containing stripped horse serum in order to avoid exposure to any lactogenic hormones present in normal serum. Lactogenic hormones were then added [I (5  $\mu$ g/ml), hydrocortisone (1  $\mu$ g/ml; Sigma), and ovine M (1  $\mu$ g/ml; NIH)] for various times. When cells were grown in the absence of I, 0.5% BSA was added as a substitute. I withdrawal has been performed in this manner and when added back to cells in the presence of hydrocortisone and M, the  $\beta$ -casein gene shows the normal induction (Poyet, P., unpublished observations). In experiments using cycloheximide, cells were treated with hormones for 48 h and then incubated for an additional 5 h with cycloheximide (10  $\mu$ g/ml) (12).

### RNA Isolation and Analysis

RNA was isolated as previously described (35). For Northern analysis RNA was transferred onto nylon membranes (Bio-trans, ICN Biomedicals, Irvine, CA) after denaturation with glyoxal (36). For slot blot analysis, RNA was applied to nitrocellulose after denaturation in formaldehyde (BA-84, Schleicher and Schuell, Keene, NH). DNA probes were synthesized by oligo labeling in the presence of [ $^{32}$ P]dCTP (37). Hybridization was carried out in the presence of 50% deionized formamide, 5  $\times$  SSC (1  $\times$  SSC = 0.15 M NaCl, 0.015 M sodium citrate), 5  $\times$  Denhardt's solution, 50 mM NaHPO<sub>4</sub> (pH 6.5), 1% sodium dodecyl sulfate (SDS), and 200  $\mu$ g/ml yeast RNA or salmon sperm DNA at 42 C. Washing of filters was performed first in 2  $\times$  SSC-0.1% SDS at room temperature for 30 min, followed by a second 30-min wash at 55 C in 0.1  $\times$  SSC-0.1% SDS. For quantification of slot blots, slots were cut out and treated with 40 mM NaOH for 15 min at 68 C, followed by the addition of acetic acid at a final concentration of 0.03 N and scintillation mixture (DuPont, Wilmington, DE; Aquasol universal LSC cocktail). The radioactivity of the samples was determined with a Beckman liquid scintillation system LS-250 (Palo Alto, CA).

### In Vitro Transcription Assay

Nuclei isolation and run-on transcription assays were performed essentially as described previously (5). Labeled RNA was hybridized to single stranded  $\beta$ -casein probes contained in M13 vectors which were immobilized on nitrocellulose filters. The mouse histone H4 clone was contained in a double stranded plasmid vector. Between 50–150  $\times$  10<sup>6</sup> cpm were synthesized in a transcription assay, and 30–40  $\times$  10<sup>6</sup> cpm were used in each hybridization reaction. For assays using  $\alpha$ -amanitin, 1  $\mu$ g/ml was added during the 20-min nuclear run-on incubation period. After the hybridization filters were washed as described and treated with 10  $\mu$ g/ml RNase-A and 100 U/ml RNase-T1. Nonspecific background was between 1–3 parts/million.

### [ $^3$ H]Uridine Pulse Assay

HC-11 cells were grown as usual and treated with hormones for 48 h. At the end of the 48-h hormone treatment period, medium was removed, and cells were rinsed several times with Hanks' Balanced Salt Solution and incubated for 30 min in 2 ml medium without serum in the presence of 1 mCi [ $^3$ H] uridine. Serum was not added in order to help deplete the UTP pool already present in the cells. After 30 min, nuclei were isolated, and RNA was extracted and analyzed.

## Acknowledgments

We wish to thank Dr. Li-Yuan Yu-Lee for her advice and critical reading of this manuscript. We also appreciate the expert artistic assistance of David Scarff and Deborah Townley.

Received July 5, 1990. Revision received August 15, 1990. Accepted August 16, 1990.

Address requests for reprints to: Dr. Jeffrey M. Rosen, Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030.

This work was supported by Grant CA-16303 from the NIH.

## REFERENCES

1. Hobbs AA, Richards DJ, Rosen JM 1982 Complex hormonal regulation of rat casein gene expression. *J Biol Chem* 257:3598-3605
2. Levine JF, Stockdale FE 1985 Cell-cell interactions promote mammary epithelial cell differentiation. *J Cell Biol* 100:1415-1422
3. Topper YJ, Freeman CS 1980 Multiple hormone interactions in the developmental biology of the mammary gland. *Physiol Rev* 60:1049-1106
4. Guyette WA, Matusik RJ, Rosen JM 1979 Prolactin mediated transcriptional and posttranscriptional regulation of casein gene expression. *Cell* 17:1013-1023
5. Eisenstein RS, Rosen JM 1988 Both cell substratum regulation and hormonal regulation of milk protein gene expression are exerted primarily at the posttranscriptional level. *Mol Cell Biol* 8:3183-3190
6. Rosen JM, Eisenstein RS, Schlein AR, Poyet P, Lee K-F, Yu-Lee, Y 1988 Hormonal and cell-substratum regulation of milk protein gene expression at the posttranscriptional level. *Prog Cancer* 35:134-141
7. Danielson KG, Oborn CJ, Durban JS, Butel JS, Medina D 1984 Epithelial mouse mammary cell line exhibiting normal morphogenesis *in vivo* and functional differentiation *in vitro*. *Proc Natl Acad Sci* 81:3756-3760
8. Bentley DL, Groudine M 1986 A block to elongation is largely responsible for decreased transcription of *c-myc* in differentiated HL60 cells 321:702-706
9. Gilmour DS, Lis JT 1986 RNA polymerase II interacts with the promoter region of the noninduced *hsp70* gene in *Drosophila melanogaster* cells. *Mol Cell Biol* 6:3984-3989
10. Reines D, Wells D, Chamberlin MJ, Kane CM 1987 Identification of intrinsic termination sites *in vitro* for RNA polymerase II within eukaryotic gene sequences. *J Mol Biol* 196:299-312
11. Yoshimura M, Oka T 1989 Isolation and structural analysis of the mouse  $\beta$ -casein gene. *Gene* 78:267-275
12. Poyet P, Henning SJ, Rosen JM 1989 Hormone-dependent  $\beta$ -casein mRNA stabilization requires ongoing protein synthesis. *Mol Endocrinol* 3:1961-1968
13. Yoshimura M, Oka T 1990 Hormonal induction of  $\beta$ -casein gene expression: requirement of ongoing protein synthesis for transcription. *Endocrinology* 126:427-433
14. Wu J, Grindlay GJ, Bushel P, Mendelsohn L, Allan M 1990 Negative regulation of the human E-globin gene by transcriptional interference: role of an ALU repetitive element. *Mol Cell Biol* 10:1209-1216
15. Kimelman D, Kirschner MW 1989 An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in *Xenopus* oocytes. *Cell* 59:687-696
16. Izant JG, Weintraub H 1984 Inhibition of thymidine kinase gene expression by antisense RNA: a molecular approach to genetic analysis. *Cell* 36:1007-1015
17. Ball RK, Friis RR, Schoenenberger CA, Doppler W, Groner B 1988 Prolactin regulation of  $\beta$ -casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary cell line. *EMBO J* 7:2089-2095
18. Chomczynski P, Qasba P, Topper YJ 1984 Essential role of insulin in transcription of the rat 25,000 molecular weight casein gene. *Science* 226:1326-1329
19. Chomczynski P, Qasba P, Topper YJ 1986 Transcriptional and post-transcriptional roles of glucocorticoid in the expression of the rat 25,000 molecular weight casein gene. *Biochem Biophys Res Commun* 134:812-818
20. Yoshimura M, Oka T 1990 Transfection of  $\beta$ -casein chimeric gene and hormonal induction of its expression in primary murine mammary epithelial cells. *Proc Natl Acad Sci USA* 87:3670-3674
21. Doppler W, Groner B, Ball RK 1989 Prolactin and glucocorticoid hormones synergistically induce expression of transfected rat  $\beta$ -casein gene promoter constructs in a mammary epithelial cell line. *Proc Natl Acad Sci USA* 86:104-108
- 21a. Höck W, Hofer P, Groner B, Ball RK 1990 Prolactin and glucocorticoid hormones control transcription of the  $\beta$ -casein gene by kinetically distinct mechanisms. *Mol Endocrinol* 4:912-919
22. Nepveu A, Marcu KB 1986 Intragenic pausing and antisense transcription within the murine *c-myc* locus *EMBO J* 5:2859-2865
23. Rougvie AE, Lis JT 1988 The RNA polymerase II molecule at the 5' end of the uninduced *hsp70* gene of *D. melanogaster* is transcriptionally engaged. *Cell* 54:795-804
24. Watson RJ 1988 A transcriptional arrest mechanism involved in controlling constitutive levels of mouse *c-myc* mRNA. *Oncogene* 2:267-272
25. Maa M-C, Chinsky JM, Ramamurthy V, Martin BD, Kellems RE 1990 Identification of transcription stop sites at the 5' and 3' ends of the murine adenosine deaminase gene. *J Biol Chem* 265:12513-12519
26. Maderious A, Chen-Kiang S 1984 Pausing and premature termination of human RNA polymerase II during transcription of adenovirus *in vivo* and *in vitro*. *Proc Natl Acad Sci USA* 81:5931-5935
27. Kao S-Y, Calman AF, Luciw PA, Peterlin BM 1987 Antitermination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* 330:489-494
28. Fort P, Rech J, Vie A, Piechaczyk M, Bonnieu A, Jeanteur P, Blanchard J-M 1987 Regulation of *c-fos* gene expression in hamster fibroblasts: initiation and elongation of transcription and mRNA degradation. *Nucleic Acids Res* 15:5657-5667
29. Konings DAM, Bredenbeek PJ, Noten JFH, Hogeweg P, Spaan WJM 1988 Differential premature termination of transcription as a proposed mechanism for the regulation of coronavirus gene expression. *Nucleic Acids Res* 16:10849-10860
30. Chinsky JM, Maa M-C, Ramamurthy V, Kellems RE 1989 Adenosine deaminase gene expression: tissue dependent regulation of transcriptional elongation. *J Biol Chem* 264:14561-14565
31. Zilz ND, Murray MB, Towle HC 1990 Identification of multiple thyroid hormone response elements located far upstream from the rat  $S_{14}$  promoter. *J Biol Chem* 265:8136-8143
32. Gupta P, Rosen JM, D'Eustachio P, Ruddle FH 1982 Localization of the casein gene family to a single mouse chromosome. *J Cell Biol* 93:199-204
33. Jones WK, Yu-Lee L, Clift SM, Brown TL, Rosen JM 1985 Rat casein multigene family: fine structure and evolution of the E1-casein gene. *J Biol Chem* 260:7042-7050
34. Messing J 1983 New M13 vectors for cloning. *Methods Enzymol* 101:20-89
35. Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ 1979 Isolation of biologically active RNA from sources enriched in ribonucleases. *Biochemistry* 18:5294-5299
36. McMaster GK, Carmichael GG 1977 Analysis of single- and double stranded nucleic acids on polyacrylamide and agarose gels using glyoxal acridine orange. *Proc Natl Acad Sci USA* 74:4835-4839
37. Feinberg AP, Vogelstein B 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 132:6-13