Variation in the polyadenyllylation site of bovine prolactin mRNA

[dpT₈-N-N'] primers/sequence analysis/cDNA clones/mRNA processing]

NANCY L. SASAVAGE*,†, MICHAEL SMITH‡, SHIRLEY GILLAM‡, RICHARD P. WOYCHIK*§, and FRITZ M. ROTTMANN*‡§¶

*Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824; and ‡Department of Biochemistry, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, V6T 1W5 Canada

Communicated by Philip Leder, August 20, 1981

ABSTRACT The poly(A) site of bovine prolactin (bPRL) mRNA was examined by phased priming of cDNA synthesis with oligodeoxynucleotides of the general sequence dp(T₈-N-N'). The existence of multiple poly(A)-adjacent sequences in bPRL mRNA was indicated by the production of specific chain-termination fragments with at least three dp(T₈-N-N') sequences. Comparison of the sequence bands produced by initiation of cDNA synthesis on the bPRL mRNA template with dp(T₈-A-G), dp(T₈-G-A), and dp(T₈-C-G) revealed a shifted pattern of identical fragments. The shift in position of related sequence bands on the gel suggested that the difference in length of the three major bPRL mRNA species occurred within a span of 12 nucleotides. Sequence analysis conducted with the three dp(T₈-N-N') primers gave identical nucleotide sequences for the 3' noncoding region of the bPRL mRNA species and suggested that the mRNA molecules were heterogeneous in length. The existence of multiple poly(A) sites was confirmed by determination of the nucleotide sequence of bPRL cDNA clones containing two of the three major poly(A)-adjacent sequences predicted by the oligodeoxynucleotide primers. The mRNA molecules containing these multiple poly(A)-addition sites were shown to be present in the bPRL mRNA obtained from a single pituitary gland. The variation in the poly(A) junction of bPRL mRNA may be a reflection of the processing events at the 3' terminus of mRNAs.

The events that constitute the synthesis and maturation of mRNA molecules are a major focus of interest in the study of gene expression (1). Such studies require the precise identification of transcription unit boundaries for specific mRNA sequences (2). We have developed a method to determine poly(A)-adjacent sequences of enriched mRNA species directly without cDNA cloning (3). The technique involves phased priming of cDNA synthesis at the poly(A) junction of the mRNA template with oligodeoxynucleotides of the general sequence dp(T₈-N-N').

This communication reports the finding of multiple poly(A)-addition sites for bovine prolactin (bPRL) mRNA by dideoxy sequence analysis with the dp(T₈-N-N') primers. The results indicate that the 3' noncoding region of bPRL mRNA is heterogeneous in length, but not in sequence, and is polyadenylated at several major and minor sites within a span of 12 nucleotides.

MATERIALS AND METHODS

Materials. T7 exonuclease was a generous gift of P. Sadowski (University of Toronto) and reverse transcriptase from avian myeloblastosis virus was provided by J. Beard (Life Sciences, St. Petersburg, FL). Exonuclease III and restriction enzymes were purchased from Bethesda Research Laboratories and used as specified. All other enzymes and materials were obtained or prepared as described (3).

Purification of bPRL mRNA. Polysomal RNA was prepared from fresh bovine anterior pituitary glands (4). Enrichment for PRL mRNA was obtained by sucrose density gradient sedimentation of poly(A)-containing RNA (4). The final PRL mRNA preparation used in this study was estimated to be 80-90% homogeneous by nitro transference (data not shown).

For the purification of PRL mRNA from a single animal, total RNA was isolated from a single pituitary gland by the method of Glisin et al. (5). PRL mRNA sequences were enriched by adsorption to PRL-specific DNA cellulose (6). The bound mRNA fraction was eluted and used as a template for cDNA synthesis as described below.

Synthesis and Use of Oligodeoxynucleotide Primers. The conditions for the enzymatic synthesis of the oligodeoxynucleotide primers dp(T₈-N) and dp(T₈-N-N') have been described in detail by Gillam and Smith (7). Screening of the oligodeoxynucleotide sequences for specific initiation of cDNA synthesis on the bPRL mRNA template was performed as described (3). Reactions using the primers for determining nucleotide sequences by the chain termination and chemical cleavage methods were as reported for bovine growth hormone (bGH; somatotropin) mRNA (3).

Screening and Sequencing of bPRL cDNA Clones. A library of bPRL-positive cDNA clones prepared by Sasavage et al. (8) was screened (9) with a nick-translated restriction fragment from a bPRL insert of known sequence (8). Several positive colonies were selected and the plasmids were prepared for sequence analysis as described (8, 10, 11). Briefly, plasmids linearized with EcoRI were digested with exonuclease III or T7 exonuclease to form single-stranded templates. The same restriction fragment was used as a primer for the DNA sequence-determination reactions by the dideoxy method (10, 12).

RESULTS

Screening of dp(T₈-N-N') Primers for Specific Initiation of bPRL cDNA Synthesis. Enriched bPRL mRNA was screened with the 12 dp(T₈-N-N') sequences to determine the two nucleotides adjacent to the poly(A) tail. Surprisingly, more than one primer sequence produced specific bands in the chain termination reaction with the bPRL mRNA template. cDNA synthesis primed by the sequences dp(T₈-C-G), dp(T₈-A-G), and dp(T₈-G-C) resulted in the isolation of cDNA clones from the bPRL mRNA library. These clones were identified by the restriction endonuclease analysis of the full-length inserts.

Abbreviations: bPRL, bovine prolactin; bGH, bovine growth hormone (somatotropin); ddTTP, 2',3'-dideoxythymidine triphosphate.

* Present address: Bethesda Research Laboratories, P.O. Box 6009, Gaithersburg, MD 20879.
† Present address: Department of Microbiology, Case Western Reserve Univ., Cleveland, OH 44106.
‡ To whom reprint requests should be addressed.
d(pT₅-G-A) resulted in intense chain termination fragments on the sequencing gel (Fig. 1). Additionally, faint bands were observed in the lanes for the primers d(pT₅-C-C) and d(pT₅-A-T). The identical pattern of the chain termination fragments indicated that all the primers initiated cDNA synthesis on the same mRNA template, and the intensities of the bands reflected the relative amounts of the different templates.

Comparison of these lanes on the gel revealed a set of characteristic fragments that was easily identifiable with each of the primer sequences. This pattern of T₅-space-T₃ was shifted in location in the sequencing gel. Because denaturing sequencing gels separate oligonucleotides that differ in length by a single nucleotide, it is possible to estimate the relative size difference of a sequence band by counting the number of positions the band has shifted. For example, we estimated that the first residue of this pattern produced with d(pT₅-A-G) was shifted seven nucleotides upward relative to its position in the lane for d(pT₅-G-A) and 12 nucleotides upward relative to d(pT₅-C-G). By this analysis, we concluded that, of the three major species, the mRNA template, and primed by d(pT₅-A-G) was the longest, followed by d(pT₅-G-A) and d(pT₅-C-G), respectively.

In the case of the primer d(pT₅-G-C), a rather strong but nonidentical set of chain termination fragments was observed (Fig. 1). We had previously determined that this sequence is complementary to the poly(A) junction of bGH mRNA (3). The bPRL mRNA utilized in the primer screening assay was enriched from pituitary polysomal mRNA (13) by fractionation in sucrose density gradients; however, the major contaminant of this preparation as determined by in vitro translation was GH mRNA (data not shown). It therefore is likely that the bands observed with the oligodeoxynucleotide d(pT₅-G-C) result from the presence of GH mRNA molecules in the enriched PRL mRNA preparation.

Sequence Analysis with d(pT₅-N-N') Primers. We chose to analyze the 3' noncoding region of bPRL mRNA further with the complementary d(pT₅-N-N') sequences that produced the most intense sequence bands, namely, d(pT₅-A-G), d(pT₅-G-A), and d(pT₅-C-G). The primers were used to initiate specifically reverse transcriptase-directed cDNA synthesis in a complete set of dideoxy sequencing reactions. Each oligodeoxynucleotide primer resulted in an identical cDNA sequence for approximately 150 nucleotides of the 3' noncoding region of bPRL mRNA. A representative sequencing gel from analysis of bPRL mRNA with the primer d(pT₅-C-G) is shown in Fig. 2. The sequence obtained by this method was positively established as that of bPRL mRNA by identification of codons for several carboxy-terminal amino acids of the bPRL protein (14) and an ochre termination codon (Fig. 3).

Further analysis of the 3' terminus of bPRL mRNA by the chemical method for DNA sequence determination (15) was
necessary to complete a small portion of the sequence. We utilized \[^{32}P\]d(pT8-N-N'), the sequence complementary to the longest mRNA template, to initiate cDNA synthesis for subsequent sequence analysis of the product. The sequence data obtained by this method (data not shown) confirmed the previous dideoxy sequence analysis and completed the sequence of nucleotides immediately adjacent to the primer that had not been obtained from the chain termination method. Examination of the nucleotide sequence from this area immediately preceding the poly(A) junction of the mRNA suggested an explanation for the initiation of cDNA synthesis by multiple d(pT8-N-N') sequences (Fig. 3, sequence C). Complementary sites were present in the bPRL mRNA template for each of the dinucleotide sequences A-G, C-A, and G-C. Furthermore, the number of nucleotides between the putative poly(A) sites was as previously estimated on the primer screening gel. Together these results predicted the existence of multiple species of bPRL mRNA that differ only in the site of polyadenylation within a span of 12 nucleotides.

Identification of bPRL cDNA Clones with Multiple Poly(A)-Adjacent Sequences. To document further the existence of multiple poly(A) sites in bPRL mRNA molecules we examined a library of bPRL cDNA clones for the heterogeneous 3' termini predicted by the primer analysis. We previously had determined the sequence of two bPRL cDNA clones (pBPRL 4 and 72) which contained the poly(A)-adjacent nucleotides complementary to d(pT8-A-G) \[^{[8]}\]. In the present analysis, PRL-positive colonies were screened with a nick-translated restriction fragment that maps in the 3' noncoding region of the sequence. DNA sequence analysis was performed by the method of Smith \[^{[10]}\] utilizing a double-stranded restriction fragment as a primer for DNA synthesis on the single-stranded plasmid templates. Among the plasmids analyzed, a bPRL insert was identified that contained the poly(A)-adjacent nucleotides complementary to the primer d(pT8-G-A).

Fig. 4 compares sequencing gels obtained with the bPRL cDNA clones containing poly(A)-adjacent sequences complementary to the d(pT8-A-G) and d(pT8-C-G) primers. The sequences shown on both gels read in the same sense as the mRNA and are identical, except that in gel A there are seven additional nucleotides preceding the poly(A) segment. The identification of two bPRL cDNA clones with different poly(A) sites confirms the existence of the heterogeneous-sized transcripts predicted by the d(pT8-N-N') primer analysis.

Multiple Poly(A) Sites of bPRL mRNA from a Single Animal. Because the enriched bPRL mRNA used in this study was isolated from several animals, we considered the possibility that the variation in the 3' terminus originated from a number of alleles in the gene pool of cattle or duplicated loci within a single animal. Therefore, PRL mRNA was purified from a single pituitary by hybridization of total cytoplasmic RNA to bPRL-specific cDNA cellulose (6). The hybridized fraction of RNA was tested with the three oligodeoxynucleotide primers, d(pT8-A-G), d(pT8-C-G), and d(pT8-G-C) for specific initiation of cDNA synthesis (Fig. 5). The pattern of fragments was identical with that found in the previous analysis of mixed bPRL mRNA, suggesting that the multiple poly(A) sites of bPRL mRNA occur within the transcripts of a single anterior pituitary gland.

**DISCUSSION**

We have screened 12 oligodeoxynucleotides of the general sequence d(pT8-N-N'), utilizing chain termination of cDNA synthesis, to determine the poly(A)-addition site of bPRL mRNA. In a previous study, sequence analysis of the 3' terminus of partially purified bGH mRNA was accomplished by this method with a single d(pT8-N-N') sequence \[^{[3]}\]. A dramatic difference was observed in the previous analysis of bPRL mRNA in which
FIG. 4. Sequencing gel autoradiographs of poly(A)-adjacent sequences in bPRL cDNA clones. bPRL cDNA clones were examined for heterogeneous poly(A)-addition sites by dideoxy sequence analysis. The DNA sequence shown at the sides reads in the same sense as the mRNA. A run of A residues in the sequence is indicated as poly(A).

(A) Cloned DNA sequence originating from a bPRL mRNA species complementary to d(pT8-A-G). (B) In contrast, the DNA sequence complementary to d(pT8-G-A) is seven nucleotides shorter and also contains the poly(A) sequence. The sequencing gels were run with an eight-lane format of nucleotides to facilitate reading of the sequence. The dark area across the top of all the lanes is a compression of G residues originating from the dG:dC tails used in the cloning.

several d(pT8-N-N') sequences functioned as specific primers of PRL cDNA synthesis (Fig. 1), suggesting a variability in the poly(A)-addition site of bPRL mRNA. The sequences d(pT8-A-G), d(pT8-G-A), and d(pT8-C-G) serve as primers to major species of bPRL mRNA as observed from the intensity of the specific bands on the primer screening gel. Two other sequences, d(pT8-C-C) and d(pT8-A-T), produced less-intense bands and possibly indicate the presence of minor species of bPRL mRNA. Attempts to titrate the mRNA population with the complementary d(pT8-N-N') sequences were consistent with the presence of varying amounts of the templates (data not shown). Furthermore, the primer screening data predict that the multiple species of bPRL mRNA occur as a variation in the site of polyadenylation within a span of 12 nucleotides.

We considered the possibility that the initiation of bPRL cDNA synthesis by more than one d(pT8-N-N') sequence was an artifact resulting from a favorable mRNA secondary structure close to the poly(A) tail. The identification of bPRL cDNA cloned sequences containing two of the three major poly(A)-addition sites provides conclusive evidence that the primer data are correct (Fig. 4). Although we have yet to identify the third major poly(A) site in a bPRL cDNA clone, the results obtained by direct dideoxy sequence analysis of the mRNA suggest that this third major bPRL mRNA species also exists (Figs. 2 and 3).

This variation in the poly(A)-addition site and in amounts of specific bPRL mRNAs of varying length may reflect the nature of the polyadenylation event in eukaryotic cells. At present, the mechanisms for termination of transcription and polyadenylation are unclear (2). It has been proposed that the hexanucleotide sequence A-A-U-A-A-A located approximately 20 nucleotides from the poly(A) tail of eukaryotic mRNAs functions as a signal for polyadenylation (16). However, the exact location of this sequence from the start of the poly(A) site is variable and ranges from 11 to 30 nucleotides. Fitzgerald and Shenk (17) examined the effects of deletion mutations in simian virus 40 DNA surrounding the hexanucleotide sequence on the polyadenylation of late mRNAs. Their results strongly indicate that the A-A-U-A-A-A sequence is indeed required for poly(A) addition and that the location of the hexanucleotide sequence influences the selection of the poly(A) site. Recent evidence also suggests that, in certain instances, transcription proceeds beyond the final poly(A) site (18–22). Our results may indicate that an endonucleolytic processing event prior to polyadenylation (20) is imprecise, thus generating multiple sites for the polyadenylation.

Although this type of ragged 3' terminus has not been previously observed for other well-characterized mRNA sequences, we have also obtained evidence for multiple poly(A) sites with chicken ovalbumin mRNA by the primer screening technique described here (data not shown). Such heterogeneities in this and other mRNAs may have been overlooked in the characterization of sequences derived from single clones. However, the variation in the poly(A)-addition site of mRNA species does not appear to be universal. For example, when we ex-
amined the poly(A) junction of bGH mRNA, we found a single 3'-terminal poly(A) sequence (3). This difference in the specificity of polyadenylylation may implicate the influence of some other signal in addition to the hexanucleotide sequence in the processing events at the 3' terminus. Presumably this signal would be downstream of the observed poly(A) junction and attenuate the specificity of the polyadenylylation site.

At this time we cannot rule out the possibility that the multiple poly(A)-addition sites of bPRL mRNA arise from nonallelic or duplicated loci for the PRL gene rather than a processing event. Fig. 5 shows that the three major poly(A) sites occur in the bPRL mRNA isolated from a single animal. Such multiple genes with different 3' flanking sequences that influence the 3'-terminal processing events could also account for the variation in polyadenylylation.

Many genomic and mRNA sequences of eukaryotic proteins are now available. These analyses have provided a basis for comparison and identification of potentially important sequences in eukaryotic gene expression. At present, however, the poly(A)-addition site of some mRNAs has only been inferred from genomic sequences (23). The use of the d(pT₈-N-N') primers to establish the poly(A)-addition site(s) of mRNA sequences may be of value in determining the mechanism of the 3'-terminal processing events.

This work was supported by U.S. Public Health Service Grant CA 13175 to F.M.R.; M. S. is a Career Investigator of the Medical Research Council of Canada. This publication is Michigan Agricultural Experiment Station Journal no. 9996.