We have constructed a double-stranded cDNA library using total poly(A)-containing RNA extracted from 8-day lactating rat mammary gland and have utilized this library to isolate clones for each of the four major milk proteins. These four cDNA clones, representing the three major rat caseins and α-lactalbumin, were initially identified by colony hybridization with labeled cDNA probes synthesized from individual mRNA fractions purified by preparative gel electrophoresis. Additional characterization was accomplished by hybridizing individual clones labeled with 32P by nick translation to a Northern gel blot of an enriched fraction of the four major milk protein mRNAs. The individual mRNAs were clearly resolved by electrophoresis on fully denaturing methylmercury hydroxide agarose gels. The identity of each milk protein clone was further established by the location of unique restriction enzyme sites within each clone. Final identification of each clone was performed by hybrid-arrested cell-free translation. The sizes of the milk protein cDNA clones ranged from 70% for the α-lactalbumin gene to essentially full length for the γ-casein gene, in comparison to their respective mRNAs. This represents the first isolation of a family of peptide hormone-responsive genes.

The mammary gland provides a unique system for studying the interactions of several steroid and peptide hormones, which are required for its development and differentiation into a specialized, secretory gland (1). It is also the first system in which a peptide hormone, prolactin, has been shown to elicit a rapid effect on gene expression (2). These studies were made possible by the development of an in vitro mammary gland organ culture system (3) and have utilized the messenger RNAs that code for the casein proteins as specific biochemical markers of hormone action. Prolactin, in the differentiated gland, induces the rapid accumulation of casein mRNA (4) and hydrocortisone potentiates this response. The caseins are a family of related milk phosphoproteins which comprise 70 to 80% of the total protein found in rat milk. There are three major rat casein proteins; their mRNAs have been previously identified and partially purified by a combination of sizing techniques and affinity chromatography (5). Casein mRNA sequences from a number of other species have also been purified by similar techniques as well as by the selective immunoprecipitation of polysomes (6-8). The fourth major milk protein, α-lactalbumin, modifies the substrate specificity of galactosyltransferase, resulting in the production of the milk sugar, lactose. α-Lactalbumin, like casein, has been shown to accumulate in the mammary gland in response to prolactin, as measured by both enzymatic activity (9) and immunoprecipitation (10) of the newly synthesized protein. α-Lactalbumin mRNA has been characterized previously (11) and complementary DNA probes synthesized from the purified mRNA used to quantitate mRNA levels during mammary development (12).

Although the casein and α-lactalbumin genes are all hormonally responsive, their expression may not be regulated by identical mechanisms. Nardacci et al. (13) have reported unequal rates of accumulation of the casein and α-lactalbumin mRNAs during pregnancy and lactation in the rat and have, therefore, concluded that these genes may be differentially regulated. Ono and Oka (14) recently have reported the differential action of cortisol (in the presence of prolactin) on the accumulation of the casein and α-lactalbumin in mouse mammary gland organ culture.

Recent studies in our laboratory have revealed multiple modes of prolactin regulation of casein mRNA accumulation in rat mammary gland organ culture (2). A 2- to 4-fold increase in the synthesis of casein specific sequences detected by pulse labeling with 3H)uridine was noted within 1 h after prolactin addition. This apparent increase in casein gene transcription, however, could not account for the observed 7-fold accumulation of casein mRNA sequences. In fact, a 17- to 25-fold increase in the half-life of casein mRNA in the presence of prolactin was also observed. In order to better understand the multiple effects of hormones on the expression of the milk protein gene family, the precise kinetics of synthesis and turnover of each of these hormonally responsive genes need to be analyzed and compared to other, constitutively synthesized, nonresponsive genes. This type of analysis requires the availability of specific hybridization probes for each gene sequence. Our previous studies have employed complementary DNA probes or cDNA affinity columns, synthesized using a mixture of the α- and β-casein mRNAs as templates. Since the four milk protein mRNAs comprise almost 80% of the mRNA population in a lactating rat mammary gland, they are readily purified from the majority of nonmilk protein mRNAs. However, it has been quite difficult to completely resolve them from each other. Even the most precise sizing techniques resulted in limited sequence cross-contamination due to co-purification of similarly sized mRNA fragments (15). With the use of recombinant DNA technology, it is possible to generate chemically unique probes for each of the milk protein sequences. This should allow more precise kinetic measurements and permit a detailed analysis of the regulation of each member of this gene family. These probes are currently being employed to study the primary transcription products and the

* This work was supported by National Institutes of Health Grant CA16303. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of a Research Career Development Award (CA-00154) from the National Institutes of Health. To whom reprint requests should be addressed. Reprints will be provided only upon receipt of a stamped, self-addressed envelope.

(Received for publication, July 7, 1980, and in revised form, August 6, 1980)

Donald A. Richards, John R. Rodgers, Scott C. Supowit, and Jeffrey M. Rosen
From the Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030

Construction and Preliminary Characterization of the Rat Casein and α-Lactalbumin cDNA Clones*

The mammary gland provides a unique system for studying the interactions of several steroid and peptide hormones, which are required for its development and differentiation into a specialized, secretory gland (1). It is also the first system in which a peptide hormone, prolactin, has been shown to elicit a rapid effect on gene expression (2). These studies were made possible by the development of an in vitro mammary gland organ culture system (3) and have utilized the messenger RNAs that code for the casein proteins as specific biochemical markers of hormone action. Prolactin, in the differentiated gland, induces the rapid accumulation of casein mRNA (4) and hydrocortisone potentiates this response.

The caseins are a family of related milk phosphoproteins which comprise 70 to 80% of the total protein found in rat milk. There are three major rat casein proteins; their mRNAs have been previously identified and partially purified by a combination of sizing techniques and affinity chromatography (5). Casein mRNA sequences from a number of other species have also been purified by similar techniques as well as by the selective immunoprecipitation of polysomes (6-8). The fourth major milk protein, α-lactalbumin, modifies the substrate specificity of galactosyltransferase, resulting in the production of the milk sugar, lactose. α-Lactalbumin, like casein, has been shown to accumulate in the mammary gland in response to prolactin, as measured by both enzymatic activity (9) and immunoprecipitation (10) of the newly synthesized protein. α-Lactalbumin mRNA has been characterized previously (11) and complementary DNA probes synthesized from the purified mRNA used to quantitate mRNA levels during mammary development (12).

Although the casein and α-lactalbumin genes are all hormonally responsive, their expression may not be regulated by identical mechanisms. Nardacci et al. (13) have reported unequal rates of accumulation of the casein and α-lactalbumin mRNAs during pregnancy and lactation in the rat and have, therefore, concluded that these genes may be differentially regulated. Ono and Oka (14) recently have reported the differential action of cortisol (in the presence of prolactin) on the accumulation of the casein and α-lactalbumin in mouse mammary gland organ culture.

Recent studies in our laboratory have revealed multiple modes of prolactin regulation of casein mRNA accumulation in rat mammary gland organ culture (2). A 2- to 4-fold increase in the synthesis of casein specific sequences detected by pulse labeling with 3H)uridine was noted within 1 h after prolactin addition. This apparent increase in casein gene transcription, however, could not account for the observed 7-fold accumulation of casein mRNA sequences. In fact, a 17- to 25-fold increase in the half-life of casein mRNA in the presence of prolactin was also observed. In order to better understand the multiple effects of hormones on the expression of the milk protein gene family, the precise kinetics of synthesis and turnover of each of these hormonally responsive genes need to be analyzed and compared to other, constitutively synthesized, nonresponsive genes. This type of analysis requires the availability of specific hybridization probes for each gene sequence. Our previous studies have employed complementary DNA probes or cDNA affinity columns, synthesized using a mixture of the α- and β-casein mRNAs as templates. Since the four milk protein mRNAs comprise almost 80% of the mRNA population in a lactating rat mammary gland, they are readily purified from the majority of nonmilk protein mRNAs. However, it has been quite difficult to completely resolve them from each other. Even the most precise sizing techniques resulted in limited sequence cross-contamination due to co-purification of similarly sized mRNA fragments (15). With the use of recombinant DNA technology, it is possible to generate chemically unique probes for each of the milk protein sequences. This should allow more precise kinetic measurements and permit a detailed analysis of the regulation of each member of this gene family. These probes are currently being employed to study the primary transcription products and the

* This work was supported by National Institutes of Health Grant CA16303. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of a Research Career Development Award (CA-00154) from the National Institutes of Health. To whom reprint requests should be addressed. Reprints will be provided only upon receipt of a stamped, self-addressed envelope.
In the first procedure, the purified Sepharose 4B fractions were electrophoresed on a 1.5% agarose-urea slab gel (20) and individual bands were excised after visualization by ethidium bromide staining. The RNA was extracted from the gel after the gel slices were treated with a 21-gauge needle and RNA eluted at room temperature in 2 volumes of 50 mM Tris-HCl, pH 7.6, 500 mM NaCl, 5 mM NaEDTA, 1% SDS. After 48 h, intact ribosomal RNA was precipitated with 2 volumes of ethanol in the presence of 0.25 M NaCl. The second method of preparation of avian myeloblastosis virus RNA-directed DNA polymerase activity has been described previously (21). The individual mRNA fractions, designated 18S (α- and β-casein mRNA), 12S (γ-casein mRNA), and 10S (α-lactalbumin mRNA), were used as templates for the synthesis of cDNA as described above except 125 μCi of [α-32P]dCTP was substituted for the [3H]dCTP.

Colony Hybridization—Three sets of replicate plates were prepared for the colony hybridizations. The lysate and subsequent hybridization was carried out by a modification of the original procedure of Grunstein and Hogness (22) as described previously (23).

Isolation of Plasmid—Individual colonies were isolated and grown in 3 ml of culture. The plasmids were purified by the procedure described by Katz et al. (24). The clear lysate was then extracted once with 1 volume of phenol saturated with 50 mM Tris-HCl, pH 7.6, and 10 mM EDTA and the supernatant precipitated overnight at -20°C by the addition of 2 volumes of ethanol. After centrifugation at 25,000 × g for 30 min, the pellet was resuspended and plasmid DNA isolated by the standard phenol-chloroform extraction. Plasmid DNA was then ethanol precipitated, either after the addition of 3 volumes of distilled water or following dialysis to remove CsCl.

"Northern" Gel Analysis—Recombinant plasmid DNA was nick-translated by a modification of the procedure of Ripy et al. (36) as described (27). The final reaction volume of 50 μl contained: 1 μg of plasmid DNA, 18 units of E. coli DNA polymerase 1 (Boehringer Mannheim), 250 μCi of [α-35P]dATP (300 Ci/mmol, New England Nuclear), 30 μM concentration each of dCTP, dGTP and dTTP, 50 mM MgCl2, pH 7.6, and 150 mM NaCl. The isolated plasmid fraction was ethanol precipitated, either after the addition of 3 volumes of distilled water or following dialysis to remove CsCl.

Milk Protein cDNA Clones

In this study, we report the construction and characterization of the three major casein cDNA clones and the fourth major prolactin-like gene, α-lactalbumin.

EXPERIMENTAL PROCEDURES

Poly(A) RNA Isolation—RNA was isolated from 8-day lactating mammary glands using a phenol/chloroform/sodium dodecyl sulfate (SDS)1 extraction procedure at pH 8.0 as described (2). Total poly(A) RNA was isolated by two passages of heat-denatured, cellular RNA over oligo(dT)-cellulose (15).

Synthesis of Double Strand DNA (dsDNA)—The self-priming ability of the cDNA permitted the synthesis of the second DNA strand. Complementary DNA was synthesized from total poly(A) RNA using the following modifications of our previously published conditions (2): 50 mM Tris-HCl, pH 8.3, 12 mM MgCl2, 20 mM dithiothreitol, 20 μg/ml of oligo(dT)12-18, 150 mM KCl, 4 mM Na2PO4, 400 μM concentration each of dATP, dTTP, and dGTP, 100 μM dCTP, 5 μM [α-32P]dCTP (30 Ci/mmol), 250 units/ml of avian myeloblastosis virus RNA-directed DNA polymerase, and 10 μg of poly(A) RNA in a final volume of 1 ml. The reaction was carried out at 46°C for 20 min and terminated by the addition of 0.2 ml of 0.1 M Na2EDTA, pH 8.0. Following Sephadex G-50 chromatography, RNA was hydrolyzed at 37°C for 2 h with 0.2 n NaOH and 10 mM Na2EDTA, the reaction was neutralized with 5 M NaOAc, pH 4.5, and then precipitated by the addition of 2 volumes of ethanol.

Synthesis of Double Strand DNA (dsDNA)—The self-priming ability of the cDNA permitted the synthesis of the second DNA strand using the following conditions: 50 mM Tris-HCl, pH 8.3, 12 mM MgCl2, 20 mM dithiothreitol, 20 μg/ml of oligo(dT)12-18, 150 mM KCl, 4 mM Na2PO4, 400 μM concentration each of dATP, dTTP, dGTP, and dCTP, 100 μM dCTP, 10 μCi of [α-32P]dCTP (30 Ci/mmol), 250 units/ml of avian myeloblastosis virus RNA-directed DNA polymerase, and 15 μg of cDNA in a final reaction mixture of 0.2 ml. The incubation was performed at 37°C for 2 h.

Digestion of dsDNA—The self-priming loop of the dsDNA was cleaved with Pst I and tailed with dG to a length of 16 bases (44). It was incubated at 37°C for 2 h.

S1 Digestion of dsDNA—The self-priming loop of the dsDNA was nicked with S1 nuclease (Miles) in a reaction mixture containing 0.6 M NaCl, 0.2 M NaOAc, pH 4.5, 2.5 mM ZnCl2, and 60 units of enzyme/ml. The average dC tail size was determined by a modification of the original procedure of Ripy et al. (36) as described (25). The isolated plasmid fraction was ethanol precipitated, either after the addition of 3 volumes of distilled water or following dialysis to remove CsCl.

Concentration of Recombinant Plasmid—The size-selected dsDNA was “tailed” with dCTP using terminal deoxynucleotidyltransferase and 1 mM CoCl2 as described previously (18). The average dC tail size was determined to be 22 bases. The dsDNA was then cleaved with Pst I and with dG to a length of 16 bases (44). It was generously supplied by Dr. Achilles Dugaiczyk (Department of Cell Biology, Baylor College of Medicine). An equimolar mixture of the tailed dsDNA and plasmid was annealed in 0.5 M NaCl, 10 mM Tris-HCl, pH 7.0, and 2 mM EDTA by first heating to 60°C, followed by slow cooling to room temperature over a 4-h period.

Transformation—All manipulations of the recombinant plasmid were carried out initially in a P3 facility using the EK2 host Escherichia coli X 1776 in accordance with the National Institutes of Health “Guidelines for Research Involving Recombinant DNA Molecules” in effect at that time. Subsequent experiments were carried out under EK2, P2 and then EK1, P1 containment in accordance with more recent guidelines. Transformation was carried out as described by Norgard et al. (19). Transformants were originally plated out on L-plates (1% Bacto-tryptone, 1% NaCl, 0.4% yeast extract, and 1% Bacto-agar) supplemented with 12.5 μg/ml of tetracycline, 100 μg/ml of diaminopimelic acid, 20 μg/ml of thymidine, 20 μg/ml of cycloserine, and 1 μg/ml of biotin. The transformation efficiency of the recombinant plasmid was 3 × 1010 colonies/μg of plasmid which was 10% of that achieved with nonrecombinant plasmid.

Preparation of Screening Probes—The individual milk protein mRNA fractions used for the synthesis of [32P]DNA screening probes were isolated from total poly(A) RNA by Sepharose 4B chromatography (15) and preparative agarose-urea gel electrophoresis. Preparative gel electrophoresis was performed by two different procedures.
and heated at 68°C for 5 min. Four micrograms of each DNA sample was added to 0.15 µg of a milk protein mRNA fraction in 25 µl containing 80% formamide, 10 µm 1,4-piperazinediethanesulfonic acid, pH 6.4, 0.3 M NaCl, and 7.5 mM EDTA, pH 6.4. The samples were incubated at 46°C for 2 h. Control samples, in which the hybrids were denatured at 68°C for 10 min, were quick frozen in acetone/dry ice. The samples were then precipitated in ethanol following addition of 1 pg of purified E. coli tRNA. Cell-free translation reactions were performed using the nuclease-treated rabbit reticulocyte lysate system according to the method of Pelham and Jackson (31). The lysate and translation reagents were supplied in a kit purchased from New England Nuclear.

RESULTS AND DISCUSSION

Construction of Double-stranded cDNA Library—The general scheme for the cloning of the four major milk protein mRNA-derived structural genes is shown in Fig. 1 and detailed under “Experimental Procedures.” Total poly(A)-containing RNA (100 µg) isolated from 8-day lactating mammary glands (Fig. 2, lane 4) was used as the template to synthesize the cDNA rather than individual purified mRNA fractions. This was advantageous because the isolation of purified mRNA fractions via preparative gel electrophoresis was a relatively inefficient procedure with yields ranging from 5 to 20%. This procedure was, therefore, suitable only for the preparation of small amounts (5 to 10 µg) of the individual purified mRNAs required for screening probes. A second advantage of cloning the total poly(A) RNA fraction is that by using this protocol it is possible to establish a “double-stranded cDNA library” (35). This “library” could then be used to isolate nonmilk protein structural genes that presumably are not prolactin regulated. Such gene sequences would be useful for comparison with the prolactin-regulated milk protein gene sequences in the study of hormone action. At the stage of mammary gland development used for mRNA isolation, milk protein sequences are highly abundant, representing approximately 70 to 80% of the total poly(A) RNA (Fig. 2, lane 4). Thus, the milk protein clones would be expected to be highly abundant and easily identifiable.

Using the poly(A) RNA fraction as a template under the conditions described under “Experimental Procedures,” a yield of cDNA of 15% was obtained. In the synthesis of the second strand, 30 to 50% of the cDNA was copied into dsDNA.
Milk Protein cDNA Clones

**Fig. 4.** Northern hybridization analysis of the milk protein clones. A, an ethidium bromide-stained profile on a 2% agarose-CH$_3$HgOH gel of a Sepharose 4B fraction of 8-day lactating mRNA which is enriched for the four major milk proteins. This fraction was used for Northern hybridization analysis after transfer to DBMC paper. B, an autoradiogram of the Northern analysis using the four nick-translated plasmids initially identified by the colony hybridization.

After the S1 treatment, isolation of the full length dsDNA was accomplished by sucrose gradient centrifugation under non-denaturing conditions. Following the tailing reaction with terminal deoxynucleotidyltransferase, the remaining mass of dsDNA was only 0.5% of the starting mass of poly(A) RNA. The “dC-tailed” dsDNA, with an average size of 840 nucleotides, was annealed at a one to one molar ratio with “dG-tailed” pBR322, resulting in a total of 2.8 μg of recombinant plasmid. Following transformation, 3.75 ng of recombinant plasmid was utilized to generate sufficient colonies for screening.

**Colony Hybridizations**—Three individual mRNA fractions were prepared for the synthesis of [32P]cDNA probes to be used in the colony hybridizations as shown in Fig. 2. The first fraction (Fig. 2, lane 1) consists of the α- and β-casein mRNAs which are observed as a 15S doublet during agarose-urea gel electrophoresis. This fraction was prepared by Sepharose 4B chromatography and has previously been shown to contain greater than 90% casein mRNA, with approximately equal amounts of the α- and casein mRNAs. Complete resolution of the α- and β-casein mRNAs could not be achieved even using preparative gel electrophoresis. Therefore, a mixture of these two mRNAs was used in the initial screening procedure. The second fraction (Fig. 2, lane 2) and the third fraction (lane 3) are the 12S γ-casein and the 10S α-lactalbumin mRNAs, respectively. These fractions were also estimated to be greater than 90% pure. Synthesis of [32P]cDNA probes to each of these mRNA fractions and cross-hybridization indicated less than 10% sequence cross-contamination.

This hybridization did not appear to be due to homology among the different mRNAs.

Typical colony hybridization results, obtained using [32P]-cDNA probes generated from these three fractions, are shown in Fig. 3 for two sets of replicate plates. Even though cross-hybridization was evident among the three probes, several individual colonies produced a more intense signal for one probe than for the other two. An example is seen in the lower left corner of replicate plate set A in Fig. 3. Two arrows mark two intense signals observed with 15S probe. The signal produced with the 12S probe to the same clone was much less intense and almost no signal was obtained with the α-lactalbumin probe. Several other colonies which appeared to generate stronger signals for one of the probes as compared to the other two are also indicated by arrows.

**“Northern” Hybridizations**—Tentative gene assignments of α- or β-casein, γ-casein, and α-lactalbumin mRNAs were obtained by the colony hybridization experiments. In order to identify a versus β-casein clones and to confirm the identity of γ-casein and α-lactalbumin clones, “Northern” hybridization analysis was carried out. Six of the “15S” colony hybridization positives and three each of the γ-casein and α-lactalbumin positives were selected for analysis. Clones were grown in 100-ml cultures and plasmids were isolated as described. Each plasmid was “nick-translated” with [α-32P]nucleotides as described under “Experimental Procedures” to a specific activity of 1 × 10⁶ cpm/μg of plasmid DNA. Each probe was then hybridized to a mixture of the four purified milk protein mRNAs resolved by electrophoresis on a 2% agarose-CH$_3$HgOH gel and transferred to DBMC paper as described. Identification of α- versus β-casein clones was possible using this method because under the conditions of gel electrophoresis in the presence of 10 mM CH$_3$HgOH complete mRNA denaturation produced greater resolution of the α- and β-casein “15S doublet” mRNAs (compare to the resolution of the 15S doublet on the partially denaturing agarose-urea gels shown in Fig. 2, lane 1).

The results obtained for four of the clones selected for “Northern” analysis are shown in Fig. 4B. Plasmids B2 and B23 which were selected as “15S doublet” positives clearly hybridized to mRNA species of different lengths. Plasmid B23 hybridized with an RNA species corresponding in length with the α-casein mRNA, while plasmid B23 hybridized with an RNA species corresponding with β-casein mRNA length. Also shown are plasmids D31 and B32 originally identified as probable γ-casein and α-lactalbumin clones, which hybridized to RNA species corresponding to the 12S γ-casein and the 10S α-lactalbumin mRNAs, respectively. In each case, the hybridization was specific for each of the four probes, suggesting the lack of extensive sequence homology among the milk protein structural genes. The intensity of signals produced by each of

---

2 J. M. Rosen, unpublished observation.
the four plasmid probes also indicated that the mRNA sequence to which each cloned sequence hybridized was a relatively abundant sequence in our mRNA fraction.

More definitive proof of the identity of the recombinant DNA sequence was required since "Northern" hybridization analysis only establishes that the cloned sequence originates from an mRNA species of the correct length. A more definitive characterization of a cloned sequence may be obtained by either sequencing the cloned DNA and comparing the predicted amino acid sequence with the known amino acid sequence of that protein (36) or by performing a translational analysis such as the hybrid-arrested translation (30). We chose the latter technique since there is almost no information available about the amino acid sequences of the rat caseins and only limited published information about the a-lactalbumin amino acid sequence (37).

Hybrid-arrested translation is performed under "R loop" conditions in which mRNA/DNA hybrids are more stable than DNA/DNA hybrids. The technique requires that the cloned sequence be of sufficient length to efficiently arrest mRNA translation. Since some of the initial cloned sequences identified were relatively small, compared to the full length mRNA (for example, the a-lactalbumin plasmid, B32, was only 30% of full length), a second colony hybridization screening was carried out to isolate clones containing larger DNA inserts prior to the hybrid-arrested translation.

Rescreening for Longer DNA Inserts—We employed 32P-labeled DNA inserts as probes during the rescreening procedure. These inserts were isolated as described in the legend to Fig. 5. Rescreening was carried out for a-casein, /3-casein, and a-lactalbumin. The initial y-casein clone was estimated to be 99% full length and, therefore, it was not necessary to rescreen additional y-casein clones. As expected in the colony hybridization utilizing the cloned DNA insert sequences as probes, no cross-hybridization was observed among the various probes. The results of the rescreening are summarized in Table I. With each probe, the number of positive colonies observed represented a high percentage of the total colonies. This high percentage of positives is consistent both with the high abundance of the casein and a-lactalbumin sequences in the initial poly(A) RNA preparation and with the "Northern" hybridization results. Therefore, the hybridization observed on Northern blots could not have been to a minor mRNA species co-migrating with the more abundant milk protein mRNAs. The relative abundance of these three clones is in excellent agreement with previously reported abundancies of the casein and a-lactalbumin mRNA sequences in the 8-day lactating rat mammary gland (5, 12).

The colonies giving the strongest signals during rescreening were now selected and 100-ml cultures were established. Plasmids isolated from each positive were analyzed for insert length on a 4% polyacrylamide gel following digestion with Pst I. Approximately 85% of the Pst I sites were regenerated using the G-C tailing protocol described. If the Pst I sites were not fully regenerated, Ava II was also used to estimate the insert size. Ava II sites were detected only in the a-lactalbumin structural gene (45). The longest insert sequences obtained for each of the four milk protein clones were analyzed by polyacrylamide gel electrophoresis in comparison to a standard Hae III digest of SV40 DNA (Fig. 5). The inserts shown in this figure were all isolated by Pst I digestion and preparative gel electrophoresis prior to analysis on the 4% polyacrylamide gel as described in the figure legend. The sizes of these inserts have been compared to the lengths of the respective milk protein mRNA sequences, estimated by CH3HgOH-agarose gel electrophoresis as listed in Table II. While only one of the clones, pCy31, is essentially full length, the majority of all four of the structural gene sequences are represented in these clones. The /3-casein clone contains greater than 90% of the full length sequence. The two shorter clones, a-casein and a-lactalbumin, contain 77% and 70% of the mRNA sequences, respectively. Table II also details unique restriction enzyme sites observed within each of these structural gene sequences. These sites were useful in confirming the identity of the clones during rescreening. Several of these sites had also been predicted by restriction enzyme analysis of double-stranded cDNA prepared from isolated mRNA fractions as reported previously (38). For example, the Hha I site in the a-casein structural gene had been suggested as a key enzyme in distinguishing the a-casein from the /3-casein sequence. A Pst I site had also been predicted for one of the 155 sequences and has since proven to be a distinguishing site for the /3-casein gene.

The Hybrid-arrested Translation—Hybrid-arrested translation was now attempted for final verification of the identity of each of the clones. The individual plasmids, cleaved with Pst I, were hybridized to an mRNA fraction enriched for the three casein and a-lactalbumin mRNAs in the presence of 80% 

![Figure 5](image)

**Table I**

Rescreening of milk protein transformants with nick translated clones

<table>
<thead>
<tr>
<th>Tentative identification</th>
<th>No. of positives</th>
<th>Total screened</th>
<th>Percentage positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>a-Casein</td>
<td>25</td>
<td>114</td>
<td>22</td>
</tr>
<tr>
<td>/3-Casein</td>
<td>23</td>
<td>136</td>
<td>17</td>
</tr>
<tr>
<td>a-Lactalbumin</td>
<td>16</td>
<td>136</td>
<td>12</td>
</tr>
</tbody>
</table>

**Table II**

Sizes and unique restriction enzyme sites in the rat milk protein cDNA clones

<table>
<thead>
<tr>
<th>Clone</th>
<th>Approximate insert size</th>
<th>Percentage of mRNA</th>
<th>Unique restriction sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCa16</td>
<td>1015; /3-casein mRNA</td>
<td>77</td>
<td>HindIII, Hha I, PsI, Hpa II</td>
</tr>
<tr>
<td>pCy31</td>
<td>1117; /3-casein mRNA</td>
<td>95</td>
<td>PsI, Hpa II, Ava II, Bam HI</td>
</tr>
<tr>
<td>pLA32</td>
<td>900; a-lactalbumin mRNA</td>
<td>99</td>
<td>PsI, Hpa II, Ava II, Bam HI</td>
</tr>
<tr>
<td>pZ6</td>
<td>70; a-lactalbumin mRNA</td>
<td>70</td>
<td>PsI, Hpa II, Ava II, Bam HI</td>
</tr>
</tbody>
</table>

* Length estimates were obtained by electrophoresis on CH3HgOH-agarose gels: a-casein mRNA = 1316; /3-casein mRNA = 1170; /3-casein mRNA = 877; and a-lactalbumin mRNA = 673 nucleotides. The length of the combined GC tail was estimated to be 36 nucleotides.

* Enzymes not cutting any of the other milk protein genes.
formamide at 46°C. An aliquot of each sample was heat denatured in the 80% formamide buffer for 10 min at 68°C prior to translation as a control. A second control translation was performed using the parent plasmid pBR322 to demonstrate the absence of nonspecific inhibition and is shown in Fig. 6, lane 1. The hybrid-arrested translation with the α-casein clone is shown in Fig. 6, lane 2. The translation of the α-casein mRNA specifying the largest casein of apparent molecular weight of 42,000 (casein I) is clearly inhibited by hybridization with this cloned DNA sequence. Translation of α-casein mRNA is regained, however, if the RNA/DNA mixture is heat denatured prior to translation (Fig. 6, lane 3). Similar results were obtained for the β-casein and γ-casein clones and their respective in vitro casein proteins II and III (Fig. 6, lanes 7 to 10). For verification of the identity of the in vitro synthesized caseins, a specific antibody was used to precipitate the in vitro translated caseins. Their electrophoretic migration is shown in Fig. 6, lane 6. Since the in vivo caseins are phosphoproteins and the 42,000-dalton casein I may be glycosylated, anomalous migration patterns of these proteins have been observed during polyacrylamide gel electrophoresis even in the presence of SDS. These in vitro translation products are also synthesized as preproteins, as each of the rat caseins has been demonstrated to contain a signal peptide sequence (39). The native caseins are, therefore, unsuitable for direct comparison with the in vitro synthesized caseins. The identity of the rat β-casein gene has also been confirmed recently by direct sequence analysis of the B23 clone and comparison with the published amino sequence of the ovine and bovine β-casein proteins (42) and the rat β-casein signal peptide sequence (39).

The hybrid-arrested translation of the smallest of the major milk proteins, α-lactalbumin, is shown in Fig. 6, lane 9. Despite the fact that the cloned sequence is only 70% of full length, there is complete arrest of the in vitro translation of α-lactalbumin mRNA. The in vivo α-lactalbumin protein is also unsuitable for comparison with the in vitro product as it is a glycoprotein which has been reported to contain multiple charge forms (40). The in vitro translation product is also a preproteins whose signal peptide has previously been characterized (41). Additional verification of the identity of the cloned α-lactalbumin structural gene sequence has been obtained by comparing its restriction enzyme map with the mobility of the newly synthesized caseins which were precipitated by specific antibodies. Lanes 2 to 5 and lanes 7 to 10 show the hybrid-arrested translation with each of the four milk protein clones under normal conditions and after heat denaturation of the hybrid prior to translation as indicated by Δ.

Fig. 6. Characterization of milk protein cDNA clones by hybrid-arrested cell-free translation. A and B contain the autoradiograms of [35S]methionine-labeled cell-free translation products electrophoresed on 10% and 12% polyacrylamide gels, respectively. Lane 1 indicates that the parent plasmid pBR322 under hybrid arrest conditions does not inhibit translation. Lane 2 demonstrates the partial amino acid sequence of the rat α-lactalbumin protein.

The predicted amino acid sequence derived from the DNA recognition sequences for several different restriction enzyme sites was in excellent agreement with the actual α-lactalbumin amino acid sequence (see Richards et al., in the accompanying manuscript).

In summary, we have constructed a double-stranded cDNA library from the total poly(A) RNA isolated from 8-day lactating rat mammary gland. Using highly enriched mRNA fractions to synthesize labeled cDNA probes, we isolated clones for each of the four major milk protein structural genes. Each clone was analyzed by the “Northern” blot technique. The results of the “Northern” analysis demonstrated that each clone hybridized to an abundant mRNA species of correct length for each of the milk protein mRNAs. This analysis also indicated that the four presumptive cDNA clones were unique sequences as cross-hybridization was not observed under the conditions of stringency of hybridization employed either in the “Northern” analysis or in subsequent rescreening using the cloned insert DNA. The unique nature of each clone was further verified by the identification of restriction enzyme sites unique to each sequence. In the case of these milk protein cDNA clones, the high abundancy of their mRNA sequences was indicated by the high percentage of “positive” colonies observed upon rescreening of the double-stranded cDNA library using the cloned insert DNA. These results were in excellent agreement with the reported frequencies of these sequences in 8-day lactating poly(A) RNA. Final confirmation of the identity of each cloned sequence was obtained by hybrid-arrested cell-free translation. Specific hybridization to each of the four milk protein mRNA sequences resulted in their arrested translation as indicated. Subsequent restriction enzyme analysis of the α-lactalbumin clone is consistent with the known amino acid sequence. The availability of each of these hormonally responsive milk protein structural genes will now permit the detailed analysis of their structure and genomic organization. It should also allow further studies of the hormonal regulation of these genes during pregnancy and lactation and in mammary gland organ culture. This study represents the first isolation and preliminary characterization of a family of peptide hormone-responsive structural genes. In the following manuscript, a more detailed comparison of these genes is presented.

Acknowledgments—We wish to express our appreciation to Drs.

---

3 J. M. Rosen and D. E. Blackburn, manuscript in preparation.
REFERENCES

3. Elias, J. J. (1957) Science 125, 842-844
21. Rosen, J. M., and Morahan, J. (1980) in Laboratory Methods Manual for Hormone Action and Molecular Endocrinology (Schneider, W. T., and O'Malley, B. W., eds), pp. 4.15-4.17, Department of Cell Biology, Baylor College of Medicine, Houston, Texas
23. Dugaiczyk, A., and Woo, S. L. C. (1980) in Laboratory Methods Manual for Hormone Action and Molecular Endocrinology (Schneider, W. T., and O'Malley, B. W., eds), pp. 5.12-5.15, Department of Cell Biology, Baylor College of Medicine, Houston, Texas