

Rapid Paper

Cloning of cDNA Sequence Coding for Bovine α_{s1} -Casein

Masatoshi MAKI, Masaya NAGAO,
Masaaki HIROSE* and Hideo CHIBA

Department of Food Science and Technology,
Kyoto University, Kyoto 606, Japan

*Department of Food Science and Nutrition,
Nara Women's University,
Nara 630, Japan

Received November 4, 1982

Double-stranded cDNAs were synthesized using total polyA-containing mRNAs isolated from the mammary gland of a lactating cow as templates. *Escherichia coli* χ 1776 was transformed with the hybrid DNA of bacterial plasmid pBR322 and cDNA constructed by the dG-dC tailing method. Plasmids containing cDNA inserts at the Pst I restriction sites were selected by resistance to tetracycline and by sensitivity to ampicillin. Those containing milk protein cDNA sequences were selected by colony hybridization using radioactive (^{32} P) single-stranded cDNAs as probes. Recombinant plasmids were isolated from those clones showing intense signals on the autoradiograph, and grouped into roughly six types by restriction endonuclease cleavage patterns of the cDNA sequences. Partial nucleotide sequences of two types of clones were determined by the method of Maxam and Gilbert. The amino acid sequence predicted from the nucleotide sequence of one clone coincided with that of α_{s1} -casein, a major component of bovine milk protein. The cloned cDNA sequence is 900 nucleotides in length, but may lack a DNA sequence coding the N-terminal region as well as the signal sequence of α_{s1} -casein.

Casein proteins are abundantly synthesized in mammary glands of lactating animals under control of multi-hormonal actions.¹⁾ While biochemical knowledge of bovine casein molecules has accumulated concerning primary structures,²⁾ biosynthesis,³⁾ phosphorylation,^{4,5)} glycosylation,⁶⁾ micelle formation^{5,7)} etc., studies on the mechanism of hormonal regulation of casein gene expression have been confined to rodents.^{8~11)} Such studies on ruminants have been limited due to the nature of the animal: size, generation time, cost of feeding and management etc.

Recent development of gene engineering,

however, has removed the limitation of species. Namely, it has become possible to isolate a specific genomic DNA sequence and to transfer the foreign gene into cultured cells so as to investigate the regulation of the gene expression.^{12,13)}

To clone a specific genomic DNA sequence, it is necessary to obtain a specific probe, usually a cDNA sequence complementary to the mRNA coding for a specific protein. Complementary DNAs coding for milk proteins have been cloned in several laboratories including ours with respect to rat,^{14~16)} mouse,^{17,18)} guinea pig¹⁹⁾ and human,²⁰⁾ but not with respect to ruminants. This paper is the first report that describes cloning of a cDNA sequence coding for a α_{s1} -casein.

MATERIALS AND METHODS

Synthesis of cDNAs and construction of hybrid DNA molecules. The mammary gland of a lactating cow was obtained from a slaughterhouse and stored at -70°C until use. RNA was extracted by the SDS/phenol method and subjected to oligo(dT)-cellulose chromatography as described previously.²¹⁾ Double-stranded cDNAs complementary to the milk protein mRNAs were synthesized by reverse transcriptase using total polyA-containing RNAs as templates, and inserted into the Pst I restriction site of pBR322 by dG-dC homopolymer extension techniques as described before.¹⁵⁾

Transformation and screening. Transformation of *E. coli* χ 1776 was performed by the method of Wahl *et al.*²²⁾ Transformants grown on L-agar plates supplemented with diaminopimelic acid, 100 $\mu\text{g}/\text{ml}$, thymidine, 40 $\mu\text{g}/\text{ml}$, and tetracycline, 10 $\mu\text{g}/\text{ml}$, were tested for sensitivity to ampicillin, 20 $\mu\text{g}/\text{ml}$, to examine the presence of a cDNA insert in the plasmid. Tetracycline resistant and ampicillin sensitive ($\text{Tc}^{\text{R}}\text{Ap}^{\text{S}}$) clones were grown on a nitrocellulose filter placed on a supplemented L-agar plate, and processed for colony hybridization as described previously.¹⁵⁾ Radioactive (^{32} P) single-stranded cDNAs used as probes were synthesized using total polyA-containing mRNAs as templates, and [α - ^{32} P]dCTP (410 Ci/mmol, Amersham) as a radioactive precursor.

All cloning experiments were carried out in accordance with the P2/EK2 containment guide-lines of the Ministry of Education, Science and Culture of Japan.

Restriction endonuclease cleavage site analysis. Isolated recombinant plasmids were digested with restriction endonucleases according to the maker's instructions (Takara

Biochemicals). Electrophoresis on either 3.5% acrylamide or 1.0% agarose slab gels was performed using TBE buffer (90 mM Tris-borate, 2.5 mM EDTA, pH 8.3) as described elsewhere.^{23,24)}

DNA sequence analysis. 5'-Ends of fragmented DNAs were labeled with T4 polynucleotide kinase (from *E. coli* B, Takara Biochemicals) and [γ -³²P]ATP (6000 Ci/mmol, Amersham) after prior dephosphorylation of the DNAs with alkaline phosphatase (PL Biochemicals) as described by Maxam and Gilbert.²⁵⁾ Labeled DNAs were cleaved with suitable restriction endonucleases to get uni-terminal labeled fragments, which were separated by electrophoresis on 5% polyacrylamide gels and then eluted from crushed gels, and the recovered DNA was subjected to base modification and cleavage as described.²⁵⁾ Electrophoresis was performed on 400 mm $\times$ 0.5 mm, 10% or 15% polyacrylamide gels in TBE buffer containing 7 M urea, at 1200 volts for 6 hr.

DNA sequence data were stored and analyzed with a microcomputer (Apple II, Apple Computer Inc.) to predict amino acid sequences and to compare them with the already determined primary structures of milk proteins²⁾ (α_{s1} -, α_{s2} -, β -, κ -casein, α -lactalbumin, β -lactoglobulin).

RESULTS AND DISCUSSION

To obtain a total clone bank of DNA sequences coding for bovine milk proteins, cDNAs were synthesized using as templates total polyA-containing mRNAs isolated from the mammary gland of a lactating cow and cloned in *E. coli* χ 1776 as described in MATERIALS AND METHODS. Using 130 ng of the hybrid DNAs, 845 distinguishable colonies resistant to tetracycline (Tc^R) were obtained. Five hundred Tc^R colonies were tested for ampicillin sensitivity (Ap^S) to examine the presence of cDNA inserts in the plasmids.

A fraction of the bacterial clones (192 $Tc^R Ap^S$ clones) was further screened for carriage of cDNA sequences complementary to abundant mRNAs of the lactating mammary gland, namely, milk protein mRNAs, by the colony hybridization technique using as probes ³²P-labeled single-stranded cDNAs complementary to the total polyA-containing mRNAs. On autoradiography after the hybridization, 79 clones (41.1%) gave intense signals, whereas 56 clones (29.2%) gave weaker signals. Since the content of each mRNA coding for milk proteins is much greater than

those for non-milk proteins (about (50%)²⁶⁾ for caseins,³⁾ about 10% for β -lactoglobulin), clones showing the intense signals are supposed to carry cDNA sequences coding for milk proteins.

Identification of a cDNA sequence can be performed in several ways as follows: (1) Hybridization assay of the cloned DNA sequence with a purified mRNA.²⁷⁾ (2) Hybridization of the mRNA and the cDNA followed by *in vitro* translation with either a hybridization-arrested assay²⁸⁾ or a hybridization-selected assay.²⁹⁾ (3) Immunological detection of the cDNA-coded protein synthesized in bacteria.²⁷⁾ (4) Nucleotide sequencing of the cDNA and comparison of the predicted amino acid sequence with primary structure of a target protein. The fourth method is quite reliable and suitable for the present case where primary structures of all major milk proteins have already been determined.²⁾

Since α_{s1} -casein is the major component sharing upto 45% of the milk protein,²⁾ the content of α_{s1} -casein cDNA clones is expected to be the largest in the clone bank. In order to clone a cDNA coding for this protein easily, each recombinant plasmid isolated from the milk protein cDNA clone bank was examined for several restriction endonuclease cleavage sites and grouped roughly into one of six types

TABLE I. RESTRICTION SITE PATTERNS OF cDNA SEQUENCES

Plasmids isolated from 24 independent clones were digested with Pst I (P) to determine sizes of cDNA inserts. Those containing cDNA inserts longer than 400 bp were digested with Acc I (A), Hind III (H), Bam HI (B) or Eco RI (E). Electrophoresis was performed on either 3.5% polyacrylamide or 1.0% agarose gels as described in MATERIALS AND METHODS.

Type	Restriction site					Clone number
	P	A	H	B	E	
I	+	-	-	-	-	3, 19, 34
II	-	+	-	-	-	12, 13, 26
III	-	-	+	-	-	9
IV	-	-	-	+	-	23
V	-	-	-	-	+	1, 4, 7, 8
VI	-	-	-	-	-	2, 6, 10, 22, 24, 27

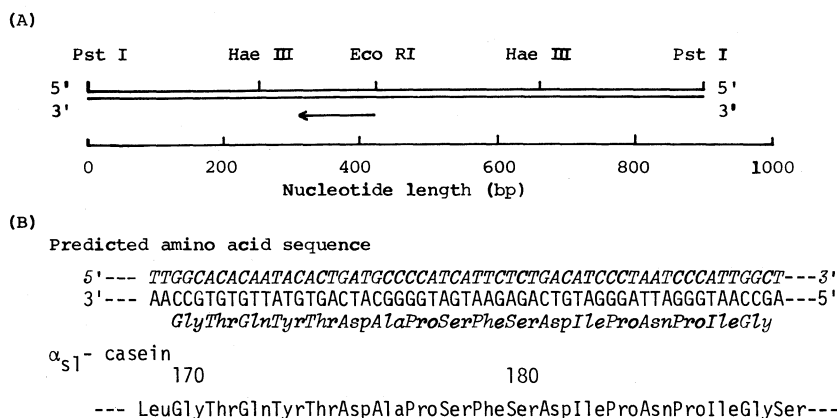


FIG. 1. The Amino Acid Sequence Predicted from the Nucleotide Sequence.

(A), an inserted cDNA (900 bp) of clone 1 was digested with Eco RI, labeled with [γ - ^{32} P]ATP, cleaved with Hae III, separated by electrophoresis on a 5% polyacrylamide gel, and sequenced by the method of Maxam and Gilbert as described in MATERIALS AND METHODS. The arrow indicates the direction of sequencing. (B), the gothic letters indicate the sequenced nucleotides, while the italicized letters indicate the complementary nucleotides. One out of six predicted amino acid sequences (italics) is aligned to be compared with the amino acid sequence of α_{s1} -casein.

by the cleavage patterns of the cDNA sequences (Table I). Two types of clones which showed abundance were selected for nucleotide sequencing (Types V and VI).

Clone 1 has a cDNA sequence of 900 nucleotides in length with an Eco RI site at its center, and one of the ^{32} P-labeled strands was partially sequenced from the restriction site (Fig. 1-A). One out of six predicted amino acid sequences (3 frames and 2 complementary strands) was consistent with the sequence of amino acids 170~187 of α_{s1} -casein²⁾ (Fig. 1-B). Thus, the cDNA of clone 1 was identified as coding for α_{s1} -casein. Considering the length of the upstream of the cDNA sequence, clone 1 appears to lack a DNA sequence of a least 200 bp for the full length including the sequence coding not only for the N-terminal region but also for the signal peptide of α_{s1} -casein.

On the other hand, amino acid sequences predicted from a partial nucleotide sequence around the Sau 3AI site of a cDNA sequence of clone 2 (about 850 bp) did not coincide with any known amino acid sequence of milk proteins (data not shown). It is still possible, however, that the sequenced region merely

does not cover a coding sequence of a milk protein.

Previously, we cloned a cDNA sequence coding for rat C2-casein,^{15,30)} a major "Ca²⁺-sensitive" casein, which is a counterpart of α_{s1} -casein of cow. First, we attempted to select an α_{s1} -casein cDNA clone from the bovine milk protein cDNA clone bank by colony hybridization using the ^{32} P-labeled C2-casein cDNA. However, no specific hybridization signal was observed on autoradiography under our conditions (data not shown). Moreover, no possible sequence of even three amino acids predicted from the partially determined nucleotide sequence of C2-casein cDNA coincided with the amino acid sequence of α_{s1} -casein. Recently, Blackburn *et al.* have reported the nucleotide sequence of rat β -casein cDNA.³¹⁾ The predicted amino acid sequence has shown that only 38% of the amino acids of β -caseins are conserved between the rat and the cow. On the other hand, reconstituted casein micelles can be formed with "Ca²⁺-sensitive" and "Ca²⁺-insensitive" casein molecules,^{5,7)} for which we have previously demonstrated that bovine and murine components are interchangeable.³²⁾ Therefore, such tremendous

divergence of the casein genes makes it interesting to further investigate the relationship of function and structure of casein molecules in micelle formation.

Acknowledgments. We wish to express our gratitude to Professor T. Komano for the generous offer of the use of the physical containment facility, and to Dr. K. Sugimoto and Dr. H. Sugisaki for their kind advice for nucleotide sequencing.

REFERENCES

- 1) H. A. Tucker, "Lactation," Vol. I, ed. by B. L. Larson and V. R. Smith, Academic Press, New York and London, 1974, p. 277.
- 2) R. M. Whitney, J. R. Brunner, K. E. Ebner, H. M. Farrell, Jr., R. V. Josephson, C. V. Morr and H. E. Swaisgood, *J. Dairy Sci.*, **59**, 795 (1976).
- 3) M. Yoshikawa, T. Mizukami, K. Omori, R. Sasaki and H. Chiba, *Agric. Biol. Chem.*, **45**, 177 (1981).
- 4) E. W. Bingham, H. M. Farrell, Jr. and J. J. Basch, *J. Biol. Chem.*, **247**, 8193 (1972).
- 5) H. Chiba and M. Yoshikawa, *Memoirs Coll. Agr.*, Kyoto Univ., No. 110, p. 1 (1977).
- 6) A. M. Fiat, C. Alais and P. Jollès, *Eur. J. Biochem.*, **27**, 408 (1972).
- 7) C. W. Slattery and R. Evard, *Biochim. Biophys. Acta*, **317**, 529 (1973).
- 8) N. M. Mehta, N. Ganguly, R. Ganguly and M. R. Banerjee, *J. Biol. Chem.*, **255**, 4430 (1980).
- 9) A. A. Hobbs, D. A. Richards, D. J. Kessler and J. M. Rosen, *J. Biol. Chem.*, **257**, 3598 (1982).
- 10) B. Teyssot, L. M. Houdebine and J. Djiane, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 6729 (1981).
- 11) M. Maki, M. Hirose and H. Chiba, *J. Biochem.*, **88**, 1845 (1980).
- 12) G. Scangos and F. H. Ruddle, *Gene*, **14**, 1 (1981).
- 13) D. M. Robins, I. Paek, P. H. Seeburg and R. Axel, *Cell*, **29**, 623 (1982).
- 14) D. A. Richards, J. R. Rodgers, S. C. Supowit and J. M. Rosen, *J. Biol. Chem.*, **256**, 526 (1981).
- 15) M. Maki, M. Ueda, M. Hirose and H. Chiba, *Agric. Biol. Chem.*, **46**, 1979 (1982).
- 16) A. M. Dandekar and P. K. Qasba, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 4853 (1981).
- 17) N. M. Mehta, M. R. El-Gewely, J. Joshi, R. B. Helling and M. R. Banerjee, *Gene*, **15**, 285 (1981).
- 18) L. G. Henninghausen and A. E. Sippel, *Eur. J. Biochem.*, **125**, 131 (1982).
- 19) L. Hall, R. K. Craig, M. R. Edbrooke and P. N. Campbell, *Nucl. Acids Res.*, **10**, 3503 (1982).
- 20) L. Hall, M. S. Davies and R. K. Craig, *Nucl. Acids Res.*, **9**, 65 (1981).
- 21) M. Maki, M. Hirose and H. Chiba, *Agric. Biol. Chem.*, **45**, 2921 (1981).
- 22) G. M. Wahl, R. A. Padgett and G. R. Stark, *J. Biol. Chem.*, **254**, 8679 (1979).
- 23) T. Maniatis, A. Jeffrey and H. van deSande, *Biochemistry*, **14**, 3787 (1975).
- 24) M. W. McDonell, M. N. Simon and F. W. Studier, *J. Mol. Biol.*, **110**, 119 (1977).
- 25) A. M. Maxam and W. Gilbert, *Methods Enzymol.*, **65**, 499 (1980).
- 26) M. Yoshikawa, T. Mizukami, R. Sasaki and H. Chiba, *Agric. Biol. Chem.*, **42**, 2185 (1978).
- 27) O. Amster, D. Salomon, O. Zemel, A. Zamir, E. P. Zeelon, F. Kantor and I. Schechter, *Nucl. Acids Res.*, **8**, 2055 (1980).
- 28) B. M. Paterson, B. E. Roberts and E. L. Kuff, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 4370 (1977).
- 29) R. P. Ricciardi, J. S. Miller and B. E. Roberts, *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 4927 (1979).
- 30) M. Hirose, T. Kato, K. Omori, M. Maki, M. Yoshikawa, R. Sasaki and H. Chiba, *Biochim. Biophys. Acta*, **667**, 309 (1981).
- 31) D. E. Blackburn, A. A. Hobbs and J. M. Rosen, *Nucl. Acids Res.*, **10**, 2295 (1982).
- 32) M. Hirose, T. Kato, K. Omori, M. Takeuchi, M. Yoshikawa, R. Sasaki and H. Chiba, *Biochim. Biophys. Acta*, **671**, 139 (1981).