Nucleotide sequence of a cDNA encoding mouse beta casein

Masami Yoshimura, Mihir R.Banerjee[1] and Takami Oka

Laboratory of Molecular and Cellular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892 and [1]Tumor Biology Laboratory, School of Life Sciences, University of Nebraska, Lincoln, NE 68588, USA

Submitted 22 September 1986

Beta casein is a major milk protein produced by the lactating mammary gland and its gene expression is hormonally controlled (1). cDNAs encoding mouse beta casein were isolated from a mouse mammary gland library prepared as described (2), using a short cDNA clone of mouse beta casein (1) as a hybridization probe. The longest cDNA clone was chosen and the DNA sequence of each strand was determined by the chain termination method (3). Templates for sequencing were prepared by either deletion (4) or Mu phage insertion (K. Mizuuchi, personal communication). The sequence of 1120 bp minus the polyA tail is presented below with the deduced amino acid sequence. The nucleotide sequence includes 55 bp of the 5'-untranslated region, the 693 bp coding region and 372 bp of the 3'-untranslated region. Comparison of our data with those reported for rat beta casein (5) revealed 86% homology at the nucleotide level and 78% homology at the amino acid level, which includes 100% homology at the 15 amino acid residues of signal peptide.

```
  1                                                          58
  ATCATCCTTTCAGCTTCACCTCCTCTCTTGTCCTCCACTAAAGGACTTGACAGCCATG
                                                        Met
59                                                          118
AAGGTCTTCATCCTCGCCTGCCTTGTGGCCCTTGCTCTTGCAAGAGAGACTACATTTACT
LysValPheIleLeuAlaCysLeuValAlaLeuAlaArgGluThrThrPheThr
119                                                         178
GTATCCTCTGAGACTGATAGTATTTCCAGTGAGGAATCTGTTGAACATATCAATGAGCAG
ValSerSerGluThrAspSerIleSerSerGluGluSerValGluHisIleAsnGluGln
179                                                         238
AAACTTCAGAAGGTGAATCTCATGGGACAGCTGCAGGCAGAGGATGTGCTCCAGGCTAAA
LysLeuGlnLysValAsnLeuMetGlyGlnLeuGlnAlaGluAspValLeuGlnAlaLys
239                                                         298
GTTCACTCCAGCATCCAGTCACAGCCCCAGGCCTTTCCATATGCTCAGGCTCAAACCATC
ValHisSerSerIleGlnSerGlnProGlnAlaPheProTyrAlaGlnAlaGlnThrIle
299                                                         358
TCTTGCAATCCCGTCCCACAAAACATCCAGCCTATTGCTCAACCCCCTGTGGTGCCATCT
SerCysAsnProValProGlnAsnIleGlnProIleAlaGlnProProValValProSer
359                                                         418
CTTGGGCCTGTCATTTCTCCTGAACTGGAATCCTTCCTTAAAGCTAAAGCCACCATCCTT
LeuGlyProValIleSerProGluLeuGluSerPheLeuLysAlaLysAlaThrIleLeu
419                                                         478
CCCAAGCACAAAACAGATGCCCCTCCTTAACTCTGAAACTGTGCTCCGCCTCATAAACTCT
ProLysHisLysGlnMetProLeuLeuAsnSerGluThrValLeuArgLeuIleAsnSer
479                                                         538
CAAATCCCCAGCCTTGCCAGTCTTGCTAATCTGCACCTTCCTCAGTCTCTGGTCCAGCTC
GlnIleProSerLeuAlaSerLeuAlaAsnLeuHisLeuProGlnSerLeuValGlnLeu
```

```
539                                                         598
CTGGCACAGGTTGTTCAGGCTTTTCCTCAGACTCACCTGGTTTCTTCTCAGACCCAGCTG
LeuAlaGlnValValGlnAlaPheProGlnThrHisLeuValSerSerGlnThrGlnLeu
599                                                         658
TCTCTTCCTCAGTCCAAAGTTCTGTACTTTCTGCAGCAAGTAGCACCCTTCCTTCCACAA
SerLeuProGlnSerLysValLeuTyrPheLeuGlnGlnValAlaProPheLeuProGln
659                                                         718
GATATGTCTGTCCAAGACCTTCTGCAGTACCTAGAACTTCTTAACCCCACCGTCCAATTC
AspMetSerValGlnAspLeuLeuGlnTyrLeuGluLeuLeuAsnProThrValGlnPhe
718                                                         778
CCTGCCACTCCACAACATTCCGTTTCTGTCTAAGAGGATTTCCAGGTTATTTTCTCCTCC
ProAlaThrProGlnHisSerValSerVal
779                                                         838
TTATGTTTTTTGAACTGACTGAAACTGGAAATGTACAAATCTTTTCACCTTTGGATCATG
839                                                         898
CTACAAAAATGATATATTTCTGAATTAATCTACATGGAAGAAAAGAAAATGTATTCTTTT
899                                                         958
ATTTATTCTATGTATTATATGGCATTCATTTGAATTCGACCCATGGACTCTATATGCTTC
959                                                        1018
TCAGTTATATTACAGGAATTTTATAAGTGTTCAATATGGAGTTGAAAATGCAAGTCAATA
1019                                                       1078
ATGTATACAAAATAGTTTGTGAAAAATTGGATTTTCTATTTTTTTTCTTTGAGAACCTAT
1079                                                       1120
TTCCTTTCCAGTCATTTCAATAAAATAATCCTTTAGGCATAT(A)n
```

REFERENCES

1. Banerjee, M.R., Ganguly, R., Mehta, N.M. and Ganguly, N. (1982) in Leung, B.S. (ed), Hormonal Regulation of Mammary Tumors, Eden Press, Montreal, Canada, Vol. II, pp. 229-283.
2. Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161-170.
3. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
4. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103-119.
5. Blackburn, D.E., Hobbs, A. and Rosen, J.M. (1982) Nucl. Acids Res. 10, 2295-2307.