 Nucleic Acids Research, Vol. 18, No. 7 1909

The cDNA and derived amino acid sequences of human and bovine bone Gla protein

Michael C.Kiefer, Andrew C.S.Saphire, Diane M.Bauer and Philip J.Barr
Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA

Submitted February 12, 1990

We have isolated and sequenced cDNAs encoding human and bovine bone Gla protein (BGP, osteocalcin), a low molecular weight polypeptide specific to bone. A λZAP cDNA library constructed from human osteosarcoma polyA+ RNA was screened with a unique ^{32}P-labeled oligonucleotide probe (27 mer) encoding the nine N-terminal amino acids of the human BGP precursor protein (1). Below is shown the nucleotide sequence of human BGP cDNA, clone hBGP-1. The sequence contains a 300 nucleotide (19–318) open reading frame encoding a 100 amino acid human BGP precursor.

A λZAP cDNA library constructed from bovine bone matrix polyA+ RNA was subsequently screened with the ^{32}P-labeled cDNA insert from hBGP-1. Below is shown the nucleotide sequence of bovine BGP cDNA, clone bBGP-3. The sequence also contains a 300 nucleotide (28–327) open reading frame encoding a 100 amino acid bovine BGP precursor.

The encoded amino aicd precursor sequences are compared below along with the previously described rat and mouse sequences (1, 2). Surprisingly, in both human and bovine sequences, an apparent alternate RNA splice site gives rise to propeptides with two additional amino acids (*) when compared to the rat and mouse sequences. Otherwise, the precursor sequences contain typical signal peptides (−51 to −29) and also propeptides (−28 to −1) that contain the information required for proteolytic maturation and γ-carboxylation of glutamic acid residues (2).

REFERENCES

1. Celeste,J. *et al.* (1986) *EMBO J.* **8**, 1885–1890.
2. Pan,L.C. and Price P.A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6109–6113.

```
  1  CGCAGCCACC GAGACACCAT GAGAGCCCTC ACACTCCTCG CCCTATTGGC CCTGGCCGCA CTTTGCATCG CTGGCCAGGC
 81  AGGTGCGAAG CCCAGCGGTG CAGAGTCCAG CAAAGGTGCA GCCTTTGTGT CCAAGCAGGA GGGCAGCGAG GTAGTGAAGA
161  GACCCAGGCG CTACCTGTAT CAATGGCTGG GAGCCCCAGT CCCCTACCCG GATCCCCTGG AGCCCAGGAG GGAGGTGTGT
241  GAGCTCAATC CGGACTGTGA CGAGTTGGCT GACCACATCG GCTTTCAGGA GGCCTATCGG CGCTTCTACG GCCCGGTCTA
321  GGGTGTCGCT CTGCTGGCCT GGCCGGCAAC CCCAGTTCTG CTCCTCTCCA GGCACCCTTC TTTCCTCTTC CCCTTGCCCT
401  TGCCCTGACC TCCCAGCCCT ATGGATGTGG GGTCCCCATC ATCCCAGCTG C
```

```
  1  GTCCACGCAG CCGCTGACAG ACACACCATG AGAACCCCCA TGCTGCTCGC CCTGCTGGCC CTGGCCACAC TCTGCCTCGC
 81  TGGCCGGGCA GATGCAAAGC CTGGTGATGC AGAGTCGGGC AAAGGCGCAG CCTTCGTGTC CAAGCAGGAG GGCAGCGAGG
161  TGGTGAAGAG ACTCAGGCGC TACCTGGACC ACTGGCTGGG AGCCCCAGCC CCCTACCCAG ATCCGCTGGA GCCCAAGAGG
241  GAGGTGTGTG AGCTCAACCC TGACTGTGAC GAGCTAGCTG ACCACATCGG CTTCCAGGAA GCCTATCGGC GCTTCTACGG
321  CCCAGTCTAG AGCTTGCAGC CCTGCCCACC TGGCTGGCAG CCCCCAGCTC TGGCTTCTCT CCAGGACCCC TCCCCTCCCC
401  GTCATCCCCG CTGCTCTAGA ATAAACTCCA GAAGAGGAAA AAAAAAAAAA AAAAAA
```

```
       -51                   -29        **                 1
Human   MRALTLLALLALAALCIAGQAGAKPSGAESSKGAAFVSKQEGSEVVKRPRRYLYQWLGAPVPYPDPLEPRREVCELNPDCDEL
Bovine  MRTPMLLALLALATLCLAGRADAKPGDAESGKGAAFVSKQEGSEVVKRLRRYLDHWLGAPAPYPDPLEPKREVCELNPDCDEL
Rat     MRTLSLLTLLALTAFCLSDLAGAKPSDSESDK--AFMSKQEGSKVVNRLRRYLNNGLGAPAPYPDPLEPHREVCELNPNCDEL
Mouse   MRTLSLLTLLALAALCLSDLTDAKPSGPESDK--AFMSKQEGNKVVNRLRRY----LGASVPSPDPLEPTREQCELNPACDEL

                        49
Human   ADHIGFQEAYRRFYGP-V
Bovine  ADHIGFQEAYRRFYGP-V
Rat     ADHIGFQDAYKRIYGTTV
Mouse   SDQYGLKTAYKRIYGITI
```