ELSEVIER

Biochimica et Biophysica Acta 1231 (1995) 117–120

BBA Biochimica et Biophysica Acta

Short Sequence-Paper

The coding sequence of the bovine ferrochelatase gene ☆

Hisashi Shibuya [a], Dan Nonneman [a], Manoel Tamassia [b], Owen L. Allphin [a], Gary S. Johnson [a,*]

[a] *Department of Veterinary Pathology, University of Missouri, 1600 East Rollins, Columbia, MO 65211, USA*
[b] *Department of Veterinary Medicine and Surgery, University of Missouri, 1600 East Rollins, Columbia, MO 65211, USA*

Received 21 April 1995; accepted 27 April 1995

Abstract

The amino acid code and surrounding regions in the bovine ferrochelatase gene were amplified by a combination of reverse transcriptase PCR and vectorette PCR and sequenced. The bovine code was 86% homologous to the human ferrochelatase code but was altered at a position corresponding to the presumed human initiator codon.

Keywords: Ferrochelatase gene; Vectorette PCR; Coding sequence; PCR; (Bovine)

Ferrochelatase occupies sites in the inner mitochondrial membrane [1] where it catalyzes the chelation of ferrous ion by protoporphyrin to complete heme biosynthesis. Protoporphyria is a disease in which protoporphyrin accumulates due to ferrochelatase deficiency. Dermal photosensitivity is a prominent symptom of this disease. Several mutations in the human ferrochelatase gene have been associated with either dominant or recessive protoporphyria [2–4].

A recessive form of protoporphyria occurs in limousin and blond de'Aquitaine cattle [5,6]. Biochemical and immunochemical analyses of residual ferrochelatase in homozygous affected cattle suggested that bovine protoporphyria results from a mutation in the coding region of the ferrochelatase gene [7,8]. In this study, we report the total coding sequence for normal bovine ferrochelatase.

Total RNA [9] from bovine blood leukocytes was used as template for routine reverse transcriptase PCR (RT/PCR). A poly-T oligonucleotide was used to prime the M-MLV reverse transcriptase. Five overlapping gene segments (Fig. 1; primers 1–10) were amplified with ferrochelatase-specific PCR primers (Table 1). Direct sequencing [10] confirmed the identity of the amplification products and revealed 1224 bp of open reading frame representing 98% of the coding region of the bovine ferrochelatase gene. However, we were unable to determine sequences at the 5′ or 3′ ends of the bovine ferrochelatase gene code by RT/PCR.

To ascertain the nucleotide sequences at the 5′ and 3′ ends of the code for bovine ferrochelatase, genomic DNA was extracted from blood leukocytes and used as template for vectorette-PCR (Vt/PCR) amplification. A standard procedure [11] was used for Vt/PCR amplification; however, our vectorette (see Table 1) was unusual in that it contained 5′ single-strand extensions at both ends. The **GATC** 5′ extension at one end of the vectorette was used to ligate to 'sticky ends' on *Mbo*l or *Bcl*1 restriction fragments. In separate experiments, the **CG** 5′ extension at the other end of the vectorette was used to ligate to 'sticky ends' on *Hpa*II restriction fragments. Orientation of the vectorette was determined by selection of the appropriate vectorette oligonucleotide for phosphorylation with T4 polynucleotide kinase prior to ligation [11].

The template for amplification of gene segments around the 3′ end of the ferrochelatase code was the vectorette ligated to *Mbo*l restriction fragments of bovine genomic DNA. The PCR amplification was done in two stages. First, vectorette-specific primer No. 20 (Table 1), was paired with bovine-ferrochelatase-specific primer #11. Although distinct bands were not detected with ethidium

☆ The sequence data reported in this paper have been submitted to the EMBL/GenBank™ Data Libraries under accession numbers L34172, L34173 and L34174.

* Corresponding author. Fax: +1 314 8822950.

SSDI 0005-2728(95)00077-1

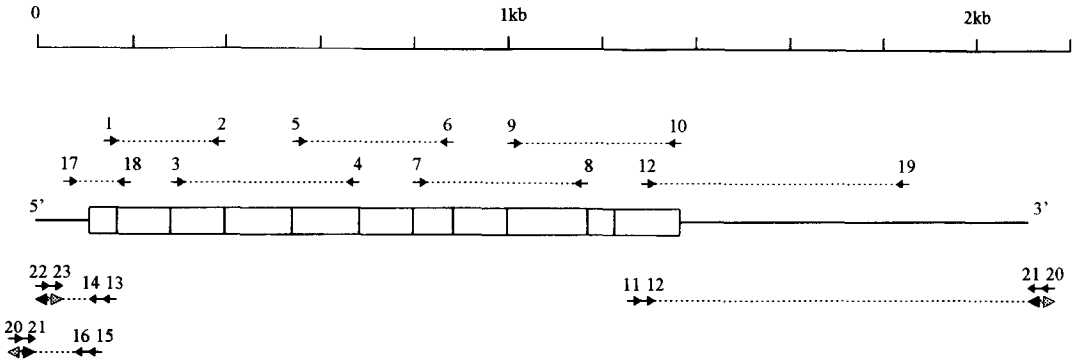

Fig. 1. The coding region of the bovine ferrochelatase gene (rectangle) is flanked by sequenced untranslated regions (lines). Dotted lines above the rectangle represent segments sequenced from RT/PCR amplicons. Dotted lines below the rectangle represent segments sequenced from Vt/PCR amplicons. Double-headed arrows show the position and orientation of the vectorette (stippled arrowhead represents the end with the 5′ **CG** overhang; solid arrowhead represents the end with the 5′ **GATC** overhang).

bromide, the product was used as template in a second stage of amplification with primers 12 and 21 (Fig. 1). This 'nested' PCR amplification produced a 0.86 kb gene fragment for direct sequencing. Because *Mbo*I restriction sites occur near the 5′ end of the bovine sequence determined by RT/PCR, the vectorette/*Mbo*I restriction fragment 'library' could not be used to study the beginning of the bovine ferrochelatase code. Therefore, the desired segment was amplified from a vectorette/*Hpa*II 'library.' In this two-stage amplification, vectorette primer No. 22 was paired with specific primer 13 and then vectorette primer 23 was used with specific primer 14. This 'nested' PCR amplification produced a 0.09 kb amplification product for direct sequencing. From this sequence, bovine ferrochelatase-specific primers 15 and 16 were designed and used consecutively with vectorette primers 20 and 21 to

Table 1
Oligonucleotide sequences

Single-copy primers:
Primer 1 5′ - CAA TGC GTT CAC TCG GCG CAA ACA TG - 3′
Primer 2 5′ - TTG GAC GGG GAG CGT CAT GAG GTC - 3′
Primer 3 5′ - GGA AGC CGA AAA CTG GAA TAT TAA TG - 3′
Primer 4 5′ - CCA CCT GTG GTG GAG CAG CTG TAC TG - 3′
Primer 5 5′ - ATA GCC CCT CAC AAA TAC TAT ATT GG - 3′
Primer 6 5′ - TTA CTT ACA GAC ATC GGC AGT GAG TG - 3′
Primer 7 5′ - CAG TGC TTT GCA GAT CAT ATT CTA AA - 3′
Primer 8 5′ - ACC TCC TTG GCT AAA ACT TGA GAG TA - 3′
Primer 9 5′ - TAG GTT GGT CCG ATG CCC TGG TTG GG - 3′
Primer 10 5′ - TCA CAG CTG TTG GCT GGT GAA GAA - 3′
Primer 11 5′ - GTC TGT CCT GCA GGC CCT GGC CGA - 3′
Primer 12 5′ - ACG CCG ATC GCG GAG CAG GAC GCC - 3′
Primer 13 5′ - GCG GAG CAG GAC GCC CGC AGA TCG - 3′
Primer 14 5′ - AAG GCC GCA GCC ATG TTT GCG CCG AC - 3′
Primer 15 5′ - CAG CCA TGT TTG CGC CGA CTG AAA GC - 3′
Primer 16 5′ - GCC GCC GGA GGA ACC CCG GTC CCG - 3′
Primer 17 5′ - TGA CAG GCC CTT GAG CCG CCG TAC AG - 3′
Primer 18 5′ - ACA CCT CCA GTC CAA GGA GCG CTG - 3′
Primer 19 5′ - TGG AGG AGC CTG GGG GCT ACA GTC CG - 3′

Vectorette primers:
Primer 20 5′ - CGC TGT CTG TAG AAG GTA AGG AAC GG - 3′
Primer 21 5′ - TCT GTA GAA GGT AAG GAA CGG ACG AG - 3′
Primer 22 5′ - TCT CGA ATC GTA AGC GTT CGT ACG AG - 3′
Primer 23 5′ - CGT TCG TAC GAG AAT CGC TGT CTC TG - 3′

Vectorette anchor*:
5′ - CGAATGCAGAGACGCTGTCTGTAGAAGGTAAGGAACGGACGAGAGAAGGGAGCA - 3′
3′ - TTACGTCTCTGTCGCTAAGAGCATGCTTGCGAATGCTAAGCTCTTCCCTCGTCTAG - 3′

[a] Underline indicates the mismatched portion of the vectorette.

atgcaccggtggggctgagtggctggttggtgctgggct 39

ccgcggctaggcggttggcggggcgggttcttggttgtgctgccgccggaggaaccccggtcccgggccgggctttcagtcggcgcaaac 129

-30

MetAlaAlaAlaLeuArgSerAlaGlyValLeuLeuArgAspArgLeuLeuTyrGlyGlySerArgAlaCysGlnProArgArgCysGln

ATGGCTGCGGCCTTGCGATCTGCGGGCGTCCTGCTCCGCGATCGGCTGCTGTACGGCGGCTCAAGGGCCTGTCAGCCACGGAGGTGCCAG 219

-1 +1

*SerGlyAlaAlaThrAlaAlaAlaAlaThrGluThrAlaGlnArgAlaArg*SerProLysProGlnAlaGlnProGlyAsnArgLysPro

TCAGGTGCGGCCACAGCGGCGGCCGCCACGGAAACCGCCCAGCGTGCCAGAAGCCCCAAGCCTCAGGCTCAGCCGGGGAACAGGAAGCCA 309

30

ArgThrGlyIleLeuMetLeuAsnMetGlyGlyProGluThrValGluGluValGlnAspPheLeuGlnArgLeuPheLeuAspGlnAsp

AGGACTGGAATATTGATGCTGAACATGGGGGGCCCGGAAACCGTCGAAGAAGTCCAGGACTTCCTTCAGAGGCTCTTTCTGGACCAAGAC 399

60

LeuMetThrLeuProValGlnAspLysLeuGlyProPheIleAlaLysArgArgThrProLysIleGlnGluGlnTyrArgArgIleGly

CTCATGACGCTCCCCGTCCAAGATAAGCTGGGACCATTCATCGCCAAACGCCGAACCCCCAAAATTCAGGAGCAGTACCGCAGGATTGGA 489

90

GlyGlySerProIleLysMetTrpThrSerLysGlnGlyGluGlyMetValLysLeuLeuAspGluLeuSerProHisThrAlaProHis

GGCGGATCGCCCATCAAGATGTGGACCTCCAAGCAGGGGGAAGGCATGGTGAAGCTGCTGGATGAGCTGTCCCCACACACAGCCCCGCAC 579

120

LysTyrTyrIleGlyPheArgTyrValHisProLeuThrGluGluAlaIleGluGluMetGluArgAspGlyLeuGluArgAlaValAla

AAATACTATATTGGATTCCGGTACGTCCACCCGTTAACAGAGGAAGCAATTGAAGAGATGGAGAGGGACGGCCTGGAGAGGGCGGTTGCT 669

150

PheThrGlnTyrProGlnTyrSerCysSerThrThrGlySerSerLeuAsnAlaIleTyrArgTyrTyrAsnGluValGlyArgLysPro

TTCACGCAGTACCCCCAGTACAGCTGCTCCACCACAGGCAGCAGCTTAAATGCCATTTACAGATACTACAATGAAGTGGGCAGGAAGCCC 759

180

ThrMetLysTrpSerThrIleAspArgTrpProThrHisProLeuLeuIleGlnCysPheAlaAspHisIleLeuLysGluLeuAspHis

ACCATGAAGTGGAGCACGATCGACAGGTGGCCCACGCACCCCCTCCTCATCCAGTGCTTTGCTGACCATATTCTGAAAGAGCTGGATCAC 849

210

PheProProGluLysArgArgGluValValIleLeuPheSerAlaHisSerLeuProMetSerValValAsnArgGlyAspProTyrPro

TTTCCTCCTGAGAAGAGACGCGAGGTGGTCATTCTTTTTTCTGCTCACTCACTGCCAATGTCTGTGGTCAACAGAGGGGACCCCTATCCT 939

240

GlnGluValGlyAlaThrValGlnArgValMetAspLysLeuGlyTyrSerAsnProTyrArgLeuValTrpGlnSerLysValGlyPro

CAAGAGGTAGGCGCCACTGTCCAGAGAGTCATGGACAAGCTGGGTTACTCCAACCCCTACCGACTTGTTTGGCAGTCCAAGGTTGGCCCG 1029

270

MetProTrpLeuGlyProGlnThrAspGluAlaIleLysGlyLeuCysLysArgGlyArgLysAsnIleLeuLeuValProIleAlaPhe

ATGCCCTGGCTGGGTCCTCAGACAGACGAGGCCATCAAAGGGCTTTGTAAGAGGGGAAGGAAGAACATCCTTTTAGTGCCAATAGCGTTT 1119

300

ThrSerAspHisIleGluThrLeuTyrGluLeuAspIleGluTyrSerGlnValLeuAlaSerGluCysGlyLeuGluAsnIleArgArg

ACCAGCGATCACATCGAAACGCTCTATGAGCTGGACATCGAGTACTCTCAAGTTCTAGCCAGTGAGTGTGGACTTGAAAACATCAGAAGA 1209

330

AlaGluSerLeuAsnGlyAsnProLeuPheSerLysAlaLeuAlaAspLeuValHisSerHisLeuGlnSerLysGluArgCysSerThr

GCGGAGTCTCTTAATGGAAACCCACTGTTCTCGAAGGCCCTGGCCGACCTGGTGCACTCACACCTCCAGTCCAAGGAGCGCTGCTCCACA 1299

360

GlnLeuThrLeuSerCysProLeuCysValAsnProThrCysArgGluThrLysSerPhePheThrSerGlnGlnLeu***

CAGCTGACTCTGAGCTGTCCGCTCTGCGTGAACCCTACCTGCAGGGAGACCAAATCCTTCTTCACCAGCCAGCAGCTGTGAccccggcgg 1389

cacgccgctgggaggtgcgcgtgcccgcctcccgacacctccgaggaggaggagggcgcatccggccgttagggaggaggttacatccgt 1479

gccggcgtagcctctggccacaaactgtgtgctttctccaggaaggggctctgaagtccttcttgagggctttctacttttatagaccaa 1569

cacagtaagcatttatggtccctcctccctgggttactagtttggaatgtccactagaccatttttaaaaatgagttttaaaaatgtatc 1659

ttctgaaatgcatacacattctcagtaccgatttgggtttggtagcaacacccccacccccccaaaaaacaaccgtgcaagggtgtgaag 1749

agtttgttttgtgggctgtcagtgtgcatggtgggggtgggggtttagtcgctcagtctgtccgactctttgcgacccacggactgtagc 1839

ccccaggctcctccatccatgggattctccaggcagagatgctggagtgggctgccacttccttctccaggcagtgtgcgtgaggtctgt 1929

ttaaagtcccgcacttgttttgggaagtgaacatagttgagtaaaggagatataattttagaaaagaaaggcgcctgctctttccctcag 2019

cccgtgggttacggaatgccggtctgatggtgtggttcccacagtgactgctcctgggaggtcactgttgatc 2092

Fig. 2. A composite nucleotide sequence for bovine ferrochelatase determined by sequencing RT/PCR and Vt/PCR amplicons. Uppercase letters show the translated region. The probable leader peptide is shown in italics. Underlined sequences were determined only from genomic DNA by Vt/PCR and have not been confirmed with RT/PCR.

```
Human      ggctggggacgcgcgtggggatcgctacccggctcggccactgctggg
Bovine     atgcaccggtggggctgagtggctggttggtgctgggctccgcggcta
Mouse                      gggcgctgtggcgcgggtgcggagcgctgcca

Human      cggacacctgggcgcgccgccgcgggaggagcccggactcgggccgag
Bovine     ggcggttggcggggcgggttcttggttgtgctgccgccggaggaaccc
Mouse      ggtcggcgttcccgggccttgccggagcgccgccgcagaggatccctg

Human      gctgcccaggcaATGCGTTCACTCGGCGCAAACATGGCTGCGGCCCTG
Bovine     cggtcccgggccgggctttcagtcggcgcaaacATGGCTGCGGCCTTG
Mouse      gccgtcccggaaATGCTTTCGGCCAGCGCCAACATGGCTGCGGCCCTG
Kozak seq.       gccrccATGG            gccrccATGG
```

Fig. 3. Nucleotide sequences surrounding possible initiator sites in the human, bovine and murine ferrochelatase gene. In-frame methionine codons and stop codons are shown in bold. The consensus sequence for initiation sites compiled by Kozak [14] is provided on the bottom line for comparison.

amplify a *Bcl*1 vectorette 'library,' extending the sequence 106 bp further upstream from the ferrochelatase code (Fig. 1).

To verify that sequences determined by Vt/PCR from genomic DNA templates were also present in cDNA, bovine-sequence primers 17 and 18 were used for RT/PCR amplification near the 5′ end of the code and bovine-sequence primers 12 and 19 were used to amplify near the 3′ end. Direct sequencing confirmed that the genomic DNA and the cDNA sequences were identical in these segments.

Fig. 2 contains a composite sequence assembled from overlapping segments obtained with RT/PCR and Vt/PCR. The bovine nucleotide code is 86% homologous to the human code [12] and 85% homologous to the mouse code [13]. A notable species difference is the **GGG** in the bovine gene at a position corresponding to that considered to be the initiator methionine codon in the human and mouse gene [12,13]. An **ATG**, 7 codons downstream, is likely to be the initiator codon for the bovine gene, as it is the first methionine downstream from an in-frame **TGA** stop codon. The human and mouse genes also have an **ATG** at a position corresponding to this bovine **ATG**. As illustrated in Fig. 3, comparisons of the sequences surrounding the two **ATG** codons in man and mouse with the consensus sequences around known initiator codons [14] suggest that the downstream **ATG**s may also initiate translation in these species.

As illustrated in Fig. 2, the probable N-terminal amino acid of mature bovine ferrochelatase, determined by analogy to the mouse gene [13], is preceded by a 47-residue leader peptide which includes eight argininyl groups and only two acidic amino acid residues. This composition is typical of a mitochondrial-targeting leader peptide [15].

References

[1] Harbin, B.M. and Daily, H.A. (1985) Biochemistry 24, 366–370.
[2] Wang, X., Poh-Fitzpatrick, M., Taketani, S., Chen, T. and Piomelli, S. (1994) Biochim. Biophys. Acta 1225, 187–190.
[3] Scheider-Yin, X., Schafer, B.W., Mohr, P., Burg, G. and Minder, E.I. (1994) Hum. Genet. 93, 711–713.
[4] Sarkany, R.P.E., Alexander, G.J.M.A. and Cox, T.M. (1994) Lancet 343, 1394–1396.
[5] Healy, P.J., Camilleri, L., Poulos, V., Hart, K.G., Kemp, B.J., Moore, R.E., Dooley, J. and Warner, M. (1992) Austral. Vet. J. 69, 144–145.
[6] Schlecher, F., Delverdier, M., Bezille, P., Cabanie, P. and Espinasse, J. (1991) Vet. Record 129, 403–407.
[7] Bloomer, J.R., Hill, H.D., Morton, K.O., Anderson-Burnham, L.A. and Straka, J.G. (1987) J. Biol. Chem. 262, 667–671.
[8] Straka, J.G., Hill, H.D., Krikava, J.M., Kools, A.M. and Bloomer, J.R. (1991) Am. J. Hum. Genet. 48, 72–78.
[9] Puissant, C. and Houdebine, L.M. (1990) BioTechniques 8, 148–149.
[10] Shibuya, H., Nonneman, D.J., Huang, T., H.-M., Ganjam, V.K., Mann, F.A. and Johnson, G.S. (1993) Anim. Genet. 24, 345–348.
[11] Edwards, A., Civitello, A., Hammond, H.A. and Caskey, C.T. (1991) Am. J. Hum. Genet. 49, 746–756.
[12] Nakahashi, Y., Taketani, S., Okuda, M., Inoue, K. and Tokunaga, R. (1990) Biochem. Biophys. Res. Commun. 173, 748–755.
[13] Taketani, S., Nakahashi, Y., Osumi, T. and Tokunaga, R. (1990) J. Biol. Chem. 265, 19377–19380.
[14] Kozak, M. (1991) J. Biol. Chem. 266, 19867–19870.
[15] Roise D. and Schatz, G. (1988) J. Biol. Chem. 263, 4509–4511.