

Sequence of human 5-aminolevulinate synthase cDNA

M.J.Bawden, I.A.Borthwick, H.M.Healy, C.P.Morris, B.K.May* and W.H.Elliott

Department of Biochemistry, University of Adelaide, Adelaide, South Australia, Australia
 Submitted September 8, 1987

Accession no.Y00451

The mitochondrial enzyme, 5-aminolevulinate synthase, catalyses the first step in the heme biosynthetic pathway and is one of the few animal genes shown to be subject to end product repression (1). We have isolated 5-aminolevulinate synthase clones from a λgt10 library of human liver cDNA by cross-hybridization with a rat 5-aminolevulinate synthase cDNA clone. A comparison with the cDNA sequences of chicken (2) and rat (unpublished) 5-aminolevulinate synthases, facilitated the assignment of the correct reading frame for the human liver 5-aminolevulinate synthase precursor. The 3' untranslated region of the mRNA contains a long poly(A) tail with an AATAAA polyadenylation signal (boxed). The predicted mature protein contains 586 amino acids and has an N-terminal transit sequence of 56 amino acids (arrowed). This sequence shows strong homology with the mature protein sequences of rat (83%) and chicken (78%), but somewhat less (45%) with that of mouse (3). Similarly, the transit sequences of human, rat, and chicken share a high degree of homology (97% and 89% respectively) in the first 38 amino acids. In contrast, the mouse transit sequence bears no homology to the other three.

```

TGCCCTTGTCGACTGAGTGCCCGCCCTCCTCGCCGGCCCTCTGCAGTCTCAGCCGAGGAGCCAGACTACTCTCTGAACATCGAGAGTGTGTGTCGCCCTGCCCAITTCATCCCGA
    val pro glu ala phe leu glu lys ala glu lys ser leu leu phe tyr ala phe tyr ala phe cys pro lys met met glu val gly ala lys phe ser met ala ser the
    val pro glu ala phe leu glu lys ala glu lys ser leu leu phe tyr ala phe tyr ala phe cys pro lys met met glu val gly ala lys phe ser met ala ser the
    GTCCCCCAGGCCCTTTCTGCAGAAAGCAACAATCTCTGTGTCTATGCCAAACTGCCCAAGATGATGGGAAGTGGGGCCAGACGCCCTCCGGATTCTCCACTCGAGCACTAGA
    150
leu pro gly asp glu glu thr pro pro ala ser glu lys asp glu lys ala lys val val glu glu thr leu met asp pro ser arg val glu met ala his ser phe arg leu asp
TACCACAAGATCAAGAAAACCCCTCCGGCCAGTGAAGAAGTCAAACCTCTAAGGCCAAGGTTCCAAAGACTCTGATGGATCCCAAGAGAGTCCAGATGGCCACACAGCTTCGGTTCGGAT
    ser val trp the pro leu ala ala thr ser glu gly thr ala ser lys cys pro phe leu ala ala ala glu met le arg glu ala ala val ser ser ala lys pro val leu ser phe arg
    TCCGTCTGGACACCCCTTGCCCAACAAGCCAGGGCACTCAAGCAAAATGCCCTTCTGGCCAGCAAGATGATCAGAGAGCCAGCAGTCTCTCTCCAAAGCCAGCTCTTGAGCTTCAGG
    270
    ser val trp the pro leu ala ala thr ser glu gly thr ala ser lys cys pro phe leu ala ala ala glu met le arg glu ala ala val ser ser ala lys pro val leu ser phe arg
    AGGTATGTGCAGAAATGCATGCCCTTAAGAAGAGGGGTCTGAAACCTCAGCAGGCCCCAGTGTGGTGTAGTGTGAAACCCGATGAAGGGAGTCCCGACTGCTGAGAAAGCTCCAG
    510
asp le met glu lys glu asp glu thr pro pro ala ser glu lys asp glu lys ala lys val val glu glu thr leu met asp pro ser arg val glu met ala his ser phe arg leu asp
GACATCATGCAAAAGCAGAGACAGAAAGATGTCTCATCTCTTCATGTAACCTCTCCAGCAAAATCTGTGTGTCCACTTTTCAGATATGAGCTGCTTTTCAGAAAGAGTGAATGAGAAAC
    asp asp his the tyr arg val phe lys the val asp arg ala his le phe pro met ala asp asp tyr ser asp ser leu le the lys lys glu val ser val trp cys ser asp
    GATGACCACACCTATCGAGTTTAAACTGTGAACCGGGCAGCACACATCTCCCATGGCAGATGACTATTCAGACTCCGCTCATCCCAAAAAGCAAGTGTCAAGTCTGGTCCAGTAAT
    750
asp tyr leu gly met ser arg his pro arg val cys gly ala val met asp the leu lys glu his gly ala gly ala gly thr arg asp le ser gly the ser lys phe his val
GACTACTAGGAATGACTGCCACCCAGCGGTGTGTGGGCAGTATGAGACACTTGAACAACATGGTCTGGGGCAGGTGGTACTAGAAATATTTCTGAACTAGTAAATCCCATGTG
    asp leu glu arg glu met asp ala lys his phe lys asp ala leu leu phe ser ser leu phe val ala his asp ser the leu phe thr leu val lys met asp pro gly cys glu
    GACTTAGCGGGAGCTGGCCAGACTCCATGGAAAGATGCCCGACTCTGTTTTCTCCCTGCTTTGTGGCCATGACTCAACCTCTCCAGCCCTGGTCCAGATGATGCCAGGCTGTGAG
    990
le tyr ser asp pro gly asp his ala ser met le glu gly le arg asp ser arg val pro lys the phe arg his asp asp val his leu asp glu leu glu thr arg ser
ATTACTCTGATTCGGGACATGCTCCCTCATGATCAAGGATCGAAAGCCCGAGTGGCCAAAGTACATCTCCGACAAATGATGTCCAGCCAGCTCAGAAACTCGTCCAAAGATCT
    asp pro ser val lys le val ala phe glu thr his pro leu glu leu leu phe glu leu cys asp val ala his leu phe val met pro lys met asp le the pro val val
    GACCCTCAGTCCCAAAGTGTGGCAITTTGAAACTGTCCATTCACTGAGTGGGGCGTGTGCTCCACTGAAAGAGCTGTGTGATGTGCCCATGATGTGGAGCAATCACCTCTCGTGGAT
    1230
    glu val his ala trp gly leu tyr gly ala arg gly gly the gly asp asp asp gly val met pro lys met asp le the ser gly the leu gly lys ala phe gly cys val gly
    GAGGTCACCCATGGGGCTTTATGGGGCTCGAGCCGAGGAGTGGGATCGGATGGAGTCATGCCAAAATGGACATCATTTCTGAAACATTGGCCAAAGCCCTTTGGTGTGTGTTGGA
    1350
cys leu le ala ser the arg ser met asp the val arg ser ty ala ala gly phe le phe thr the ser leu pro pro met leu pro ser ala leu phe val arg le
GGGTACATGCCAGCAGATGCTCTGATGGACACCGTACGGTCTATCTGCTGGCTTCATCTGCAGCATGCTGTGCCACCATGCTGGTGGCTGGAGCTGGAGTCTGTGGCGATC
    leu lys ser ala glu gly arg val arg arg glu his glu arg asp val lys leu met arg glu met leu met asp ala gly leu pro val val his cys pro ser his le le pro
    CTGAAGGCGCTGAGGCAAGGTCGTTGCGCCAGCAGCACGGCAACCTCAAACTCAGACATGAGAGATGTAATGGATGGCCGCGCCCTCCCTGTGGTCCACTGCCCAAGCCACATCATCCCT
    1590
val arg val ala asp ala ala lys asp the glu val cys asp glu leu cys asp glu leu met ser arg his asp le tyr val glu ala le asp tyr the val pro val pro ser gly glu leu leu arg
GTGCGGGTTGCAGATGCTGCTAAAACACAGAGTCTGTGATGAACTAATGAGCAGACATAAATCTACGTCGAAGCAATCAATACCTACCGTGCCCGCCGGGAAAGAGCTCCTACGG
    le ala pro the pro glu his the pro glu met met asp tyr phe glu asp leu leu val the trp lys glu val gly leu leu leu phe val pro val his ser asp ala glu cys asp phe
    ATGGCCCCACCCCTCACCACACCCAGATGATGACTCTCTTGGAAATCTGCTAGTACATGGAAGCAAGTGGGGCTGGAACCTGAAGCTCATTTCTCAGCTGAGTGCACACTC
    1830
cys arg ser leu his leu phe glu val met ser glu arg his leu ser ty phe ser gly leu ser lys leu val ser ala lys ala
TCAAGAGGCGACTGCATTTGAAGTGAATGAGTGAAGAGAGAGTCTTATTTCTCAGGCTTGAAGAGATGTGTATCTGTAAGGCTTGAGCATGACCTCAATATTTACGTTAAACCCCA
    GCCCATTCATATCCAGATGGTCTCAAGTGTCTTATATGTGAATTAAGTTATATAAATTTAATCTATGATAAAAACATAGTCTCGGATAAACTGCTGCTTAATGGTG(A).
    
```

*To whom correspondence should be addressed

REFERENCES

1. May, B.K. *et al.*, (1986) *Current Topics in Cellular Regulation* Vol. 28, 233-262.
2. Borthwick, I.A. *et al.* (1985) *Eur. J. Biochem.* **150**, 481-484.
3. Schoenhaut, D.S. and Curtis, P.J. (1986) *Gene* **48**, 55-63.