Volume 14 Number 24 1986 Nucleic Acids Research

# Nucleotide sequence of rat liver δ-aminolevulinic acid dehydratase cDNA

Terry Rogers Bishop, Laurence P.Frelin and Samuel H.Boyer

Howard Hughes Medical Institute and Division of Medical Genetics, Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

Submitted 19 November 1986

The isolation of an immunologically reactive λgt11 clone encoding δ-aminolevulinic acid dehydratase (ALAD; EC 4.2.1.24), the second enzyme of the heme biosynthetic pathway, has been described (1). The original clone, ALAD 1 (73 to 747), was used to rescreen the liver cDNA library from which it was derived and ALAD 3 (-48 to 412) and ALAD 7 (433 to 1071) were isolated. Together these three clones span the entire coding region plus 48 and 78 nucleotides of the 5' and 3' untranslated regions, respectively. Both strands were sequenced using the chain termination method. When the coding region is compared to the human ALAD coding region sequence (2), 88% of the amino acids are shared with 87% nucleotide homology. The underlined nucleotides show a perfect match to the Kozak consensus sequence for eukaryotic translation initiation (3). The lysine at the active site (4) is marked with an asterisk and the presumptive polyadenylation site is boxed.

```
                                                        -48                 -31                                      -1
                                                        GGC CTA GAC TCC GTC TTA GCA GAC TGC TGT GTT CCG CCG GCC CCC ACC
                                     30                                      60                                      90
ATG CAC CAC CAG TCC GTT CTG CAC AGT GGC TAC TTT CAC CCA CTG CTT CGG GCC TGG CAG ACC ACC CCC TCC ACC GTC AGT GCC ACC AAC
Met His His Gln Ser Val Leu His Ser Gly Tyr Phe His Pro Leu Leu Arg Ala Trp Gln Thr Thr Pro Ser Thr Val Ser Ala Thr Asn
                                    120                                     150                                     180
CTC ATC TAT CCC ATC TTT GTC ACG GAT GTT CCT GAT GAT GTC CAG CCC ATT GCC AGC CTC CCG GGA GTG GCC AGG TAT GGC GTA AAC CAG
Leu Ile Tyr Pro Ile Phe Val Thr Asp Val Pro Asp Asp Val Gln Pro Ile Ala Ser Leu Pro Gly Val Ala Arg Tyr Gly Val Asn Gln
                                    210                                     240                                     270
CTA GAG GAG ATG CTG AGA CCT CTG GTG GAA GCC GGC CTT CGC TGT GTC CTG ATC TTT GGC GTC CCC AGC AGA GTT CCC AAG GAT GAA CAG
Leu Glu Glu Met Leu Arg Pro Leu Val Glu Ala Gly Leu Arg Cys Val Leu Ile Phe Gly Val Pro Ser Arg Val Pro Lys Asp Glu Gln
                                    300                                     330                                     360
GGC TCT GCA GCT GAC TCA GAG GAC TCC CCA ACT ATT GAG GCT GTC CGT CTG CTG AGG AAG ACC TTC CCT ACC CTC CTA GTG GCC TGT GAT
Gly Ser Ala Ala Asp Ser Glu Asp Ser Pro Thr Ile Glu Ala Val Arg Leu Leu Arg Lys Thr Phe Pro Thr Leu Leu Val Ala Cys Asp
                                    390                                     420                                     450
GTC TGC CTG TGC CCC TAC ACC TCC CAT GGC CAC TGT GGC CTC CTG AGT GAA AAT GGA GCC TTC CTA GCA GAG GAA AGC CGA CAA CGG TTG
Val Cys Leu Cys Pro Tyr Thr Ser His Gly His Cys Gly Leu Leu Ser Glu Asn Gly Ala Phe Leu Ala Glu Glu Ser Arg Gln Arg Leu
                                    480                                     510                                     540
GCA GAG GTG GCA CTG GCC TAT GCC AAG GCA GGT TGT CAG GTT GTA GCT CCA TCA GAC ATG ATG GAT GGA CGA GTT GAA GCC ATC AAG GCT
Ala Glu Val Ala Leu Ala Tyr Ala Lys Ala Gly Cys Gln Val Val Ala Pro Ser Asp Met Met Asp Gly Arg Val Glu Ala Ile Lys Ala
                                    570                                     600                                     630
GCC CTG CTA AAA CAT GGA CTT GGC AAC AGG GTC TCT GTG ATG AGC TAT AGT GCC AAA TTT GCC TCC TGT TTC TAT GGT CCT TTC CGG GAT
Ala Leu Leu Lys His Gly Leu Gly Asn Arg Val Ser Val Met Ser Tyr Ser Ala Lys Phe Ala Ser Cys Phe Tyr Gly Pro Phe Arg Asp
                                    660                                     690                                     720
GCA GCT CAG TCA AGC CCA GCT TTC GGA GAC CGA CGC TGT TAT CAG CTG CCT CCT GGA GCC CGT GGC CTG GCC CTC CGA GCA GTG GCC CGG
Ala Ala Gln Ser Ser Pro Ala Phe Gly Asp Arg Arg Cys Tyr Gln Leu Pro Pro Gly Ala Arg Gly Leu Ala Leu Arg Ala Val Ala Arg
                                    750      *                              780                                     810
GAC ATT CAA GAG GGA GCT GAC ATA CTC ATG GTA AAG CCG GGA TTA CCC TAC CTG GAT ATG GTG CAG GAG GTG AAG GAC AAG CAC CCC GAG
Asp Ile Gln Glu Gly Ala Asp Ile Leu Met Val Lys Pro Gly Leu Pro Tyr Leu Asp Met Val Gln Glu Val Lys Asp Lys His Pro Glu
                                    840                                     870                                     900
CTC CCC CTC GCA GTA TAC CAG GTG TCC GGA GAG TTT GCC ATG TTG TGG CAC GGA GCC AAG GCT GGG GCC TTT GAT CTC AGG ACT GCT GTA
Leu Pro Leu Ala Val Tyr Gln Val Ser Gly Glu Phe Ala Met Leu Trp His Gly Ala Lys Ala Gly Ala Phe Asp Leu Arg Thr Ala Val
                                    930                                     960                                     990
CTG GAG TCC ATG ACG GCC TTC CGC AGA GCC GGT GCT GAC ATT ATC ATC ACC TAC TTT GCA CCC CAG CTG TTG AAG TGG CTG AAG GAA GAG
Leu Glu Ser Met Thr Ala Phe Arg Arg Ala Gly Ala Asp Ile Ile Ile Thr Tyr Phe Ala Pro Gln Leu Leu Lys Trp Leu Lys Glu Glu
                                   1020                                    1050                        1071
TGA AGG AAA GTG AAG GAC TTG GTC TTG ACA GAG CTC CCT GGG CCT TAC AGA AGG GAA AAA GTA AAG GCG CTG TTA GAA CCG
End
```

1. Bishop, T. Rogers, Cohen, P.J., Boyer, S.H., Noyes, A.N. and Frelin, L.P. (1986) Proc. Natl. Acad. Sci. USA 83:5568-5572.
2. Wetmur, J.G., Bishop, D.F., Cantelino, C. and Desnick, R.J. (1986) Proc. Natl. Acad. Sci USA 83:7703-7707.
3. Kozak, M. (1984) Nucl. Acids Res. 12(2):857-872.
4. Gibbs, P.N.B. and Jordan, P.M. (1986) Biochem. J. 236:447-451.