

Cloning and sequence of mouse erythroid  $\delta$ -aminolevulinatase dehydratase cDNA

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

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

Submitted January 4, 1989

EMBL accession no. X13752

A cDNA library in  $\lambda$ gt11 was constructed from mouse (BALB/c) reticulocyte poly(A)<sup>+</sup> RNA using oligo(dT) and reverse transcriptase for the first strand synthesis and RNase H and DNA polymerase I for second strand synthesis (1). cDNA was methylated, ligated to EcoRI linkers and exhaustively digested with EcoRI (2). Double-stranded cDNA larger than 800 bp by agarose gel electrophoretic fractionation was then ligated into dephosphorylated  $\lambda$ gt11 arms (Stratagene) and packaged (Gigapack; Stratagene) using *E. coli* Y1088 (3). The library was screened using rat liver  $\delta$ -aminolevulinatase dehydratase cDNA (4) and both strands of the longest clone (1092 base pairs), B1, was subjected to dideoxynucleotide sequencing analysis, the results of which are shown below. There are only 9 (out of 330) amino acid differences between rat liver and mouse erythroid ALA-D. The nucleotide sequence of the translated region of B1 is 96% homologous with rat  $\delta$ -aminolevulinatase dehydratase cDNA (5) and 97% of the predicted amino acids are identical.

The data reported below have the accession number X13752 in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases.

```

                                     cccgac-1
ATGCCA10CCAGTCT20TCGCACAGCG30CTACT40CCACTGCT50TCGGAGCTG60CACTGCTGCT70CCACCGT80CTGCTCA90ACTCAT100
GCATCT110TCACAGCGAT120CTCTGAT130CCAGCCT140TCGCCAG150CCAGGAG160CCAGGT170CCCTAA180CCGCTAG190AGAG200ATGCTGAG210
TCTGG220AGCTG230CTGCT240TCCT250CTCC260AGT270GGAT280GGCTG290CTGCT300AGCT310CTGAG320CTGCT330
ACTAT340TGG350CTGCT360CTGCT370CTGCT380CTGCT390CTGCT400CTGCT410CTGCT420CTGCT430CTGCT440CTGCT450CTGCT460
TCTT470GAG480CTAA490TGG500ACT510CTCC520AGG530AGC540ACAGC550CGCT560TCG570CAAG580GG590CTGCT600CTGCT610
GTGAG620CTAA630TGG640AGC650CTCA660CTGCT670CTGCT680CTGCT690CTGCT700CTGCT710CTGCT720CTGCT730CTGCT740
GCCT750CTGCT760CTGCT770CTGCT780CTGCT790CTGCT800CTGCT810CTGCT820CTGCT830CTGCT840CTGCT850CTGCT860
CCCT870CCAGC880ACT890CGCC900AGC910CA920TCC930AGG940AGCT950CA960TGCT970CA980TGG990CA1000CCG1010AT1020CC1030
GC1040CC1050CC1060CT1070CC1080CC1090CT1100CA1110GT1120CT1130CA1140GT1150CT1160CA1170GT1180CT1190CA1200GT1210CT1220
CT1230GG1240AG1250CT1260CA1270CC1280CG1290CA1300CC1310CT1320CA1330CT1340CA1350CT1360CA1370CT1380CA1390CT1400
GT1410CT1420CG1430ACT1440CT1450CT1460CT1470CT1480CT1490CT1500CT1510CT1520CT1530CT1540CT1550CT1560CT1570CT1580CT1590CT1600

```

REFERENCES

- Gubler, U. and Hoffman, B.J. (1983) *Gene* 25:263-269.
- Schwarzbauer, J.E., Tamkun, J.W., Lemischka, I.R. and Hynes, R.O. (1983) *Cell* 35, 421-431.
- Young, R.A. and Davis, R.W. (1983) *Science* 222:778-782.
- Bishop, T.R., Cohen, P.J., Boyer, S.H., Noyes, A.N. and Frelin, L.P. (1986) *Proc. Natl. Acad. Sci. USA* 83:5568-5572.
- Bishop, T.R., Frelin, L.P. and Boyer, S.H. (1986) *Nucl. Acids Res.* 14:10115.