Vol. 156, No. 1, 1988 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 14, 1988 Pages 576-580

THE COMPLETE NUCLEOTIDE SEQUENCE OF HUMAN LIVER CYTOCHROME b_5 mRNA

Min Yoo and Alan W. Steggles

Department of Biochemistry and Molecular Pathology, Northeastern Ohio Universities College of Medicine, P.O. Box 95, Rootstown, OH 44272

Received August 11, 1988

We have isolated and sequenced a cDNA clone corresponding to human liver cytochrome b_5 mRNA. The 760 base pair (bp) sequence contains the complete coding and 3' non-translated regions plus 52 bp of 5' non-translated sequence. The derived amino acid sequence showed that the previous assignment of several amino acids was in error. In addition, the sequence of the previously unknown COOH hydrophobic region has been obtained.

The amino acid sequences of the soluble portions of both the human liver and erythrocyte cytochrome b_5's have been available for some time (1,2). The two proteins show the same sequence except at positions 98 which are thr and pro for the liver and erythrocyte proteins respectively. In the former the thr is followed by a hydrophobic region of unknown length or sequence, whereas in the later the pro is the terminal amino acid (1,2). Based on data from other species, the hydrophobic region of cytochrome b_5 should contain about 36 amino acids (3-7), and function as the membrane binding domain for the liver protein.

There are three intriguing questions that relate to human liver cytochrome b_5 and its corresponding mRNA. First, what is the hydrophobic sequence of human liver cytochrome b_5? Second, what is the relationship between the liver and the erythrocyte proteins? Are they derived from two mRNAs from two different genes, or from

0006-291X/88 $1.50

two mRNAs from one gene by some alternative mRNA processing mechanism? Last, what is the molecular defect in the cytochrome b_5 gene(s) that leads to one of the forms of inherited methemoglobinemia (8). Any attempts to study these questions must start with the isolation and characterization of a suitable cDNA probe corresponding to either of the liver or erythrocyte cytochrome b_5 mRNA's.

MATERIALS AND METHODS

The human liver λgt11 cDNA libraries were obtained from Clontech Ltd. (Palo Alto, CA) and Dr. S.L.C. Woo (Baylor, TX). Enzymes were purchased from Bethesda Research Labs (Bethesda, MD), Pharmacia (Piscataway, NJ) and U.S. Biochemicals (Cleveland, OH). Radioisotopes were from ICN Ltd. (Irvine, CA). Nitrocellulose filters were purchased from Schleicher and Schuell (Keene, NH) and Millipore (Bedford, MA), and nylon membranes from New England Nuclear (Boston, MA). Low melt agarose was from FMC Co. (Rockland, ME). The plasmid vector "Bluescript" was obtained from Stratagene (San Diego, CA). All other reagents and chemicals were of the best quality available.

The λgt11 phage from the Clontech library were used to infect Y1088 cells (9) prior to plating at a density of ca. 15000 pfu/ 15 cm petri dish. Nitrocellulose filter lifts were made, amplified, denatured and baked (10,11). After prehybridization (10), the filters were placed into hybridization solution (10) containing 2 x 10^6 cpm/filter of [^{32}P] labelled DNA probe (12).

The probe was the EcoRI - DraI insert isolated from the plasmid pR1 and contains the complete coding sequence for rabbit liver cytochrome b_5 mRNA (13). After an overnight hybridization at 37°C, the filters were washed once with 6 x SSC, 0.1% SDS at room temperature for 10', then, once with 1 x SSC, 0.1% SDS for 30' at 37°C and once with 0.2xSSC, 0.1% SDS for 30' at 42°C. The filters were air dried and used to expose Kodak XAR5 film at -70°C. One Quanta III intensifying screen was used. In the initial screen of ca. 500,000 plaques, five positive signals were identified, only one was subsequently plaque purified. The insert from this phage λHa was subcloned into the EcoR I cut dephosphorylated Bluescript vector. This insert (Ha) was subsequently shown to be a partial copy of human liver cytochrome b_5 mRNA. Ha was radiolabelled (12) and then used as a probe to screen the second cDNA library at high stringency (10). From this screening of ca. 500,000 phage, twelve positive signals were identified, and all were subsequently plaque purified (10). Two phage ($\lambda Hb_{5.1}$ and $\lambda Hb_{5.2}$) were studied in detail. The inserts were subcloned, and mapped with restriction enzymes (10,14). Selected DNA fragments were isolated from low melt agarose gels and again subcloned (14).

The complete nucleotide sequences of $Hb_{5.1}$ and $Hb_{5.2}$ were determined, using both double and single stranded DNA sequencing protocols (15,16).

RESULTS AND DISCUSSION

The restriction maps of pH1 and pH2 were used to determine the DNA sequencing strategy (Fig. 1). The nucleotide sequence (Fig. 2) of pH2 contains 52, 402 and 308 bp in the 5', coding and 3' regions, respectively. The derived amino acid sequence (Fig. 2) is the same as the published sequence and confirms the reassignments at six positions (1,2). Four of these are the initial met and the three adjacent amino acids; the other two are gln for glu at position 18, and asn for asp at position 62. In addition there is a transposition of the lys and arg at positions 89 and 91. The latter were probably due to protein sequencing errors (17). The amino acids at positions 89 and 91 now agree with the majority of the published sequences (3-7,18). It will be interesting to see if the equivalent monkey sequence of lys-pro-arg (19) is changed to arg-pro-lys when the monkey liver cytochrome b_5 cDNA nucleotide sequence is obtained. From the nucleotide sequence we have obtained the previously unknown hydrophobic COOH region. This thirty-six amino acid sequence is similar to those described for rabbit, cow, pig, rat and horse (3-7). The trp trp thr asn trp sequence at positions 109-113 is conserved (3-7), except for the chicken (18). The chicken liver cytochrome b_5 sequence is also unusual in that it contains six extra amino acids at the NH_2 terminus, and one less amino acid at the COOH terminus (18). It is

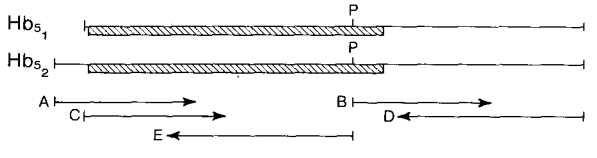

Figure 1. The sequencing strategy for human liver cytochrome b_5 cDNA. A, B, and C were double stranded, D and E were single stranded DNA sequencing. All sequences were repeated at least four times.

```
CAGCCAGCTCGACGGGGCTGTGTGTGCTGGGCCTGGCTCGCGGCGAACCGAG  52

ATG GCA GAG CAG TCG GAC GAG GCC GTG AAG TAC TAC ACC   91
Met Ala Glu Gln Ser Asp Glu Ala Val Lys Tyr Tyr Thr  (13)

CTA GAG GAG ATT CAG AAG CAC AAC CAC AGC AAG AGC ACC  130
Leu Glu Glu Ile Gln Lys His Asn His Ser Lys Ser Thr  (26)

TGG CTG ATC CTG CAC CAC AAG GTG TAC GAT TTG ACC AAA  169
Trp Leu Ile Leu His His Lys Val Tyr Asp Leu Thr Lys  (39)

TTT CTG GAA GAG CAT CCT GGT GGG GAA GAA GTT TTA AGG  208
Phe Leu Glu Glu His Pro Gly Gly Glu Glu Val Leu Arg  (52)

GAA CAA GCT GGA GGT GAC GCT ACT GAG AAC TTT GAG GAT  247
Glu Gln Ala Gly Gly Asp Ala Thr Glu Asn Phe Glu Asp  (65)

GTC GGG CAC TCT ACA GAT GCC AGG GAA ATG TCC AAA ACA  286
Val Gly His Ser Thr Asp Ala Arg Glu Met Ser Lys Thr  (78)

TTC ATC ATT GGG GAG CTC CAT CCA GAT GAC AGA CCA AAG  325
Phe Ile Ile Gly Glu Leu His Pro Asp Asp Arg Pro Lys  (91)

TTA AAC AAG CCT CCG GAA ACT CTT ATC ACT ACT ATT GAT  364
Leu Asn Lys Pro Pro Glu Thr Leu Ile Thr Thr Ile Asp (104)

TCT AGT TCC AGT TGG TGG ACC AAC TGG GTG ATC CCT GCC  403
Ser Ser Ser Ser Trp Trp Thr Asn Trp Val Ile Pro Ala (117)

ATC TCT GCA GTG GCC GTC GCC TTG ATG TAT CGC CTA TAC  442
Ile Ser Ala Val Ala Val Ala Leu Met Tyr Arg Leu Tyr (130)

ATG GCA GAG GAC TGA ACACCTCCTCAGAAGTCAGCGCAGGAAGAGC  488
Met Ala Glu Asp xxx                                 (134)

CTGCTTTGGACACGGGAGAAAAGAAGCCATTGCTAACTACTTCAACTGACA  539

GAAACCTTCACTTGAAAACAATGATTTTAATATATCTCTTTCTTTTTCTTC  590

CGACATTAGAAACAAAACAAAAAGAACTGTCCTTTCTGCGCTCAAATTTTT  641

CGAGTGTGCCTTTTTATTCATCTACTTTATTTTGATGTTTCCTTAATGTGT  692

AATTTACTTATTATAAGCATGATCTTTTAAAAATATATTTGGCTTTTAAAG  743

TAAAAAAAAAAAAAAAAAA                                  762
```

Figure 2. The complete nucleotide sequence of human liver cytochrome b_5 cDNA and the derived amino acid sequence.

possible that these differences will be restricted to avian cytochrome b5's.

We do not know whether the cDNA clone pH_2 is missing either a stretch of 3' poly A sequence, some 5' non-translated sequence, or both. The 5' non-translated sequence has the consensus ribosome binding site at nucleotides 20-30 (20), in agreement with the majority of eucaryote mRNAs.

The isolation and characterization of a cDNA for human liver cytochrome b_5 mRNA will enable studies on the one gene-two gene

question and on the molecular basis of certain inherited methemoglobinemias to be initiated (8).

ACKNOWLEDGEMENTS

This study was supported by the State of Ohio Academic and Research Challenge Funds and by a grant from the Ohio Heart Association.

REFERENCES

1. Ozols J. (1972) J. Biol. Chem. 247, 2242-2254.
2. Abe K., Kimura S., Kizawa R., Anan F.K., and Sugita Y. (1985) J. Biochem. 97, 1659-1668.
3. Takahagi Y., Gerber G.E., Nihei K. and Khorana H.A. (1980) J. Biol. Chem. 255, 1536-1541.
4. Fleming P.J., Dailey H.A., Corcoran D. and Strittmatter P. (1978) J. Biol. Chem. 253, 5369-5372.
5. Ozols J. and Gerard G. (1977) Proc. Natl. Acad. Sci. (USA) 74, 3725-3729.
6. Ozols J. and Heinemann F.S. (1982) Biochim Biophys Acta 704, 163-172.
7. Ozols J. and Gerard C. (1977) J. Biol. Chem. 252, 8549-8553.
8. Hegesh E., Hegesh J. and Kaftory A. (1986) N. Engl. J. Med. 314, 757-761.
9. Young R.A. and Davis R.W. (1983) Science 222, 778-782.
10. Maniatis T., Fritsch E.F. and Sambrook J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor.
11. Woo, SLC (1979) Methods in Enzymol. 68, 389-395.
12. Feinberg A. and Vogelstein B. Anal Biochem (1984) 137, 266-267.
13. Dariush N., Fisher C.W. and Steggles A.W. Prot. Seq. Data Analysis (1988) In press.
14. Struhl K. (1985) Biotechniques 3, 452-453.
15. Sanger F. and Coulson A.R. (1975) J. Mol. Biol. 94, 441-448.
16. Chen E.Y. and Seeberg P.H. (1985) DNA 4, 165-170.
17. Naoki A., Mayumi Y., Yoshitaka G. and Taro O. (1988) Biochem. Biophys. Res. Commun. 152, 1395-1400.
18. Zhang H. and Somerville C. (1988) Arch. Biochem. Biophys. 264, 343-347.
19. Nobrega F.A. and Ozols T. (1971) J. Biol. Chem. 246, 1706-1717.
20. Kozak M. (1987) Nucl. Acids Res. 15, 8125-8132.