The cDNA and protein sequences of human lactate dehydrogenase B

Ikuya SAKAI, Farida S. SHARIEF, Yu-Ching E. PAN* and Steven S.-L. LI†

Laboratory of Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, U.S.A.

Human lactate dehydrogenase B (LDH-B) cDNA was isolated and sequenced. The LDH-B cDNA insert consists of the protein-coding sequence (999 bp), the 5' (54 bp) and 3' (203 bp) non-coding regions, and the poly(A) tail (50 bp). The predicted sequence of 333 amino acid residues was confirmed by amino acid composition and/or sequence analyses of a total of 185 (56%) residues from tryptic peptides of human LDH-B protein. The nucleotide and amino acid sequences of the human LDH-B coding region show 68% and 75% homologies respectively with those of the human LDH-A. The peptide map and amino acid composition data have been deposited as Supplementary Publication SUP 50139 (7 pages) at the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies are available on prepayment [see Biochem. J. (1987) 241, 5].

INTRODUCTION

In mammals, the five isoenzymes of tetrameric lactate dehydrogenase (LDH) are found in various proportions among different somatic tissues and are produced *in vivo* by combination of the A (muscle) and B (heart) subunits, whereas the homotetrameric LDH-C₄ is present only in mature testis and spermatozoa [1]. The LDH-A₄, LDH-B₄ and LDH-C₄ isoenzymes possess distinct physical, catalytic and immunological properties [2,3]. The LDH-A, -B and -C polypeptides are encoded by three different genes that appear to have originated from an ancestral gene during the course of evolution [4]. In order to study the structural and evolutionary relationship and to elucidate the molecular mechanism of the tissue-specific expression of these three LDH genes, we have undertaken the investigation of protein structure and gene organization of mammalian LDH-A₄, LDH-B₄ and LDH-C₄ isoenzymes. We have reported the primary structure of human and mouse LDH-A₄ as well as mouse and rat LDH-C₄ isoenzymes [5–9]. We have also described the nucleotide sequences of human LDH-A and mouse LDH-C cDNAs of nearly full length as well as the genomic organization of human and mouse LDH-A genes and pseudogenes [7–15]. In the present work the complete amino acid sequence of human LDH-B₄ isoenzyme was determined by structural analyses of both cDNA and protein.

MATERIALS AND METHODS

Isolation and characterization of cDNA clones

Human LDH-B cDNA clones were isolated from B-lymphocyte and T-cell cDNA libraries (Clontech) by using a human LDH-A cDNA probe [8]. The *Eco*RI DNA fragments of positive clones were subcloned into M13 mp10 bacteriophages, and their partial nucleotide

sequences were determined by the dideoxy chain-termination method [16–18]. An *Eco*RI fragment of length 1.3 kb from clone hB619 was isolated from M13 double-stranded DNA and further cleaved with *Sau*3AI or *Aha*I, the DNA subfragments were cloned into M13 mp10 or M13 mp11 bacteriophages, and their nucleotide sequences were determined.

Composition and sequence analyses of tryptic peptides

The LDH-B₄ isoenzyme purified from human erythrocytes (Sigma Chemical Co.) was oxidized and digested with trypsin. Tryptic peptides were purified by two-dimensional paper chromatography and high-voltage electrophoresis at pH 4.4. The amino acid compositions of the isolated peptides were obtained by using a Dionex D500 amino acid analyser, and partial amino acid sequences of high-yield peptides were determined automatically by using a Beckman 890C protein/peptide sequencer and h.p.l.c. as previously described [5,9].

RESULTS

LDH-B cDNA sequence

The human B-lymphocyte λ gt10 cDNA library was screened by using a human LDH-A cDNA probe, and several weakly positive clones, presumably containing LDH-B cDNA, were identified. The partial nucleotide sequence of a 1.0 kb *Eco*RI fragment from clone hB537 (Fig. 1) was found to be different from that of human LDH-A cDNA, and its deduced amino acid sequence was consistent with the amino acid compositions of some tryptic peptides purified from human LDH-B protein. Thus clone hB537 contains a partial human LDH-B cDNA sequence starting from codon no. 65.

The partial human LDH-B cDNA of clone hB537 was subsequently used as a probe to identify several strongly

Abbreviation used: LDH, lactate dehydrogenase.

* Present address: Department of Protein Biochemistry, Hoffmann–La Roche, Nutley, NJ 07110, U.S.A.

† To whom correspondence and reprint requests should be addressed.

These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession no. Y00711.

Fig. 1. Restriction-endonuclease map and sequencing strategy of human LDH-B cDNA clones

The cDNA inserts were flanked by *EcoRI* (R) sites. The restriction sites given are only those used in nucleotide sequencing (*AluI*, A; *Sau3A*, S). The direction and length of each sequencing run are indicated by arrows. All nucleotide sequences for both complementary strands of the LDH-B cDNA were determined experimentally. The protein-coding sequence is shown as a solid black box. The 5' and 3' non-coding regions are indicated by open boxes, with the poly(A) tail hatched.

Fig. 2. Nucleotide and amino acid sequences of human LDH-B

The deduced amino acid sequence of human LDH-B cDNA is given below its nucleotide sequence. The results of composition and sequence analyses of tryptic peptides are denoted by hyphens and arrows respectively. The polyadenylation signal AATAAA [20] is underlined.

positive clones from the human T-cell λ gt10 cDNA library [19]. Clones hB609 and hB619 were found to contain *Eco*RI inserts of 0.9 kb and 1.3 kb respectively (Fig. 1). The complete sequence of 1307 nucleotide residues from clone hB619 (Fig. 2) consists of a single 333-amino acid-residue open reading frame between the first ATG codon triplet of translation initiation and the termination triplet TGA, the 5' (54 nucleotides) and 3' (203 nucleotides) non-coding regions and the poly(A) tail (50 nucleotides).

Characterization of tryptic peptides

The tryptic peptides of oxidized human LDH-B protein were purified by two-dimensional paper chromatography and electrophoresis at pH 4.4. The peptide map and amino acid composition data of tryptic peptides were submitted for editorial review, and are deposited as Supplementary Publication SUP 50139 at the British Library Lending Division. The results of compositional analysis as well as partial sequencing of the purified tryptic peptides were summarized in Fig. 2. A total of 185 amino acid residues obtained from tryptic-peptide analyses confirmed 56% of the amino acid sequence deduced from the LDH-B cDNA sequence.

DISCUSSION

The clone hB537 isolated from the human B-lymphocyte λ gt10 cDNA library by using a human LDH-A cDNA probe was found to contain partial LDH-B

cDNA flanked by *Pst*I sites right next to *Eco*RI linkers. One of these *Pst*I sites corresponded to codon no. 65–66 (Fig. 2). This appeared to indicate that the λ gt10 cDNA library (Clontech HL1002B) was derived from the *Pst*I fragments of its related pBR322 cDNA library (HL1002A). Thus it would be very unlikely to isolate the full-length LDH-B cDNA containing two *Pst*I sites.

The partial LDH-B cDNA of clone hB537 was subsequently used as a probe to isolate clones hB609 and hB619 from the human T-cell λ gt10 cDNA library [19]. The sequence of 1307 nucleotide residues from clone hB619 (Fig. 2) contains the full-length protein-coding regions, the 5' and 3' non-coding regions as well as poly(A) tail. The primary structure of human LDH-B protein has a calculated molecular mass of 36507 Da, excluding the initiation methionine and *N*-terminal blocked group, and its composition of 333 amino acid residues is consistent with the total amino acid analysis of human LDH-B protein (Table 1). The amino acid sequence of human LDH-B protein possesses 94% homology with the pig LDH-B sequence determined by direct protein analysis [21]. The nucleotide and amino acid sequences of human LDH-B coding region, when compared with those of human LDH-A (Fig. 3), show 68% and 75% homologies respectively [8]. It may also be noted that only 49% of the 322 nucleotide differences resulted in amino acid substitutions. The functional significance of these amino acid differences between human LDH-B₄ and LDH-A₄ isoenzymes remains to be elucidated through the use of site-directed mutagenesis.

The 3' non-coding region of human LDH-B mRNA contains a single polyadenylation signal AAUAAA [20] located at 18 nucleotide residues 5' to the poly(A) addition site. This region is much shorter than that of human LDH-A (565 nucleotide residues), and there was no obvious sequence homology between them [8]. Recently the ϵ -crystallin of duck lenses was found unexpectedly to be active LDH-B₄ isoenzyme [22]. The availability of nearly full-length human LDH-B cDNA should also facilitate the structural analysis of mammalian LDH-B genes and the molecular characterization of genetic mutations affecting the LDH-B₄ isoenzymes [23].

We thank Dr. B. H. Judd for his interest in and support of this project, Dr. C. Teng and Dr. T. Kunkel for reading the manuscript, and Ms. N. Gore and Ms. L. Roberts for typing it.

Table 1. Amino acid composition of human LDH-B protein

The values obtained from amino acid analysis were calculated from the average values of 24 h, 48 h and 72 h hydrolyses, except those for serine and threonine, which were extrapolated to zero-time hydrolysis, and those for isoleucine and valine, which were those of 72 h hydrolysis. Asparagine and aspartic acid were obtained as aspartic acid, and glutamine and glutamic acid as glutamic acid. Cysteine was determined as cysteic acid. Tryptophan was not determined.

Amino acid residue	Deduced from cDNA sequence	From amino acid analysis	
Ala	21	21.61	
Arg	8	8.37	
Asn	17	36	34.88
Asp	19		
Cys	5	32	33.24
Gln	11		
Glu	21		
Gly	23	32	22.65
His	7		
Ile	24	36	21.27
Leu	36		
Lys	26	36	26.14
Met	9		
Phe	5	36	8.36
Pro	11		
Ser	26	36	12.97
Thr	13		
Trp	6	36	13.19
Tyr	7		
Val	38	36	6.41
Total	333		

REFERENCES

- Markert, C. L., Shakelee, J. B. & Whitt, G. S. (1975) *Science* **189**, 102–114
- Holbrook, J. J., Liljas, A., Steindel, S. J. & Rossmann, M. G. (1975) *Enzymes* 3rd Ed. **11**, 191–292
- Goldberg, E. (1977) *Isozymes Curr. Top. Biol. Med. Res.* **1**, 79–124
- Li, S. S.-L., Fitch, W. M., Pan, Y.-C. E. & Sharief, F. S. (1983) *J. Biol. Chem.* **258**, 7029–7032
- Pan, Y.-C. E., Sharief, F. S., Okabe, M., Huang, S. & Li, S. S.-L. (1983) *J. Biol. Chem.* **258**, 7005–7016
- Li, S. S.-L., Feldmann, R. J., Okabe, M. & Pan, Y.-C. E. (1983) *J. Biol. Chem.* **258**, 7017–7028
- Akai, K., Yagi, H., Tian, H. F., Pan, Y.-C. E., Shimizu, M., Fong, K., Jungmann, R. A. & Li, S. S.-L. (1985) *Int. J. Biochem.* **17**, 645–648

Fig. 3. Comparison of the nucleotide and amino acid sequences of human LDH-B with those of human LDH-A

Only those nucleotides and amino acids of human LDH-A [9] that differ from those of human LDH-B are indicated. Human LDH-B polypeptide chain contains 333 amino acid residues, excluding the initiation methionine residue, and human LDH-A protein possesses 331 residues with two deletions located at positions 18 and 332 of the LDH-B sequence.

- 8 Tsujibo, H., Tiano, H. F. & Li, S. S.-L. (1985) *Eur. J. Biochem.* **147**, 9-15
9. Li, S. S.-L., Tiano, H. F., Fukasawa, K. M., Yagi, K., Shimizu, M., Sharief, F. S., Nakashima, Y. & Pan, Y.-C. E. (1985) *Eur. J. Biochem.* **149**, 215-225
10. Sakai, I., Sharief, F. S. & Li, S. S.-L. (1987) *Biochem. J.* **242**, 619-622
11. Chung, F.-Z., Tsujibo, H., Bhattacharyya, U., Sharief, F. S. & Li, S. S.-L. (1985) *Biochem. J.* **231**, 537-541
12. Fukasawa, K. M. & Li, S. S.-L. (1986) *Biochem. J.* **235**, 435-439
13. Fukasawa, K. M., Li, W.-H., Yagi, K., Luo, C.-C. & Li, S. S.-L. (1986) *Mol. Biol. Evol.* **3**, 330-342
14. Fukasawa, K. M., Tanimura, M., Sharief, F. S., Chung, F.-Z. & Li, S. S.-L. (1987) *Genetics* **115**, 177-184
15. Fukasawa, K. M. & Li, S. S.-L. (1987) *Genetics* **116**, 99-105
16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor
17. Messing, J., Crea, R. & Seeberg, P. H. (1981) *Nucleic Acids Res.* **9**, 309-321
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467
19. Yanagi, Y., Chan, A., Chin, B., Minden, M. & Mak, T. W. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3430-3434
20. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211-214
21. Kiltz, H.-H., Keil, W., Greisbach, M., Petry, K. & Meyer, H. (1977) *Hoppe-Seyler's Z. Physiol. Chem* **358**, 123-127
22. Wistow, G. J., Mulders, J. W. M. & de Jong, W. W. (1987) *Nature (London)* **326**, 622-624
23. Kanno, T., Sudo, K., Kitamura, M., Miwa, S., Ichyama, A. & Nishimura, Y. (1983) *Isozymes Curr. Top. Biol. Med. Res.* **7**, 131-150