Biochem. J. (1985) 231, 537–541 (Printed in Great Britain)

Genomic organization of human lactate dehydrogenase-A gene

Fu-Zon CHUNG, Hiroshi TSUJIBO, Uma BHATTACHARYYA, Farida S. SHARIEF 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.

A human genomic clone containing the lactate dehydrogenase-A (LDH-A) gene of approx. 12 kilobases in length was isolated and characterized. The protein-coding sequence is interrupted by six introns, and the positions of these introns are at the random coil regions or near the ends of secondary structures located on the surface of the LDH-A molecule. An additional intron is present at 24 nucleotides 5′ to the translation initiation codon ATG, while the 3′ untranslated sequence of 565 nucleotides is not interrupted. The genomic blot analysis of human placenta DNA indicates the presence of multiple LDH-A gene-related sequences.

INTRODUCTION

In mammals and birds, the three genes coding for LDH isoenzymes A_4 (muscle), B_4 (heart) and C_4 (testis) are believed to have originated from an ancestral gene during the course of evolution (Markert *et al.*, 1975; Li *et al.*, 1983*a*). The expression of these three LDH genes is developmentally regulated and tissue-specific. Further, the level of LDH-A mRNA was shown to increase following the stimulation of cultured cells by dibutyryl cyclic AMP, isoproterenol or epidermal growth factor (Miles *et al.*, 1981; Matrisian *et al.*, 1985). Recently, the LDH-A_4 protein was found unexpectedly to bind single-stranded DNA (Williams *et al.*, 1984). In order to illustrate the structural and evolutionary relationships of the LDH-A, B and C genes and to study the molecular mechanism of gene regulation, we have undertaken the investigation of protein structure and genomic organization of mammalian lactate dehydrogenase isoenzymes. We have reported the primary structure of LDH-A_4 isoenzymes from human and mouse (Tsujibo *et al.*, 1985; Li *et al.*, 1985) as well as LDH-C_4 isoenzymes from mouse and rat (Pan *et al.*, 1983; Li *et al.*, 1983*b*). We have also described the isolation and characterization of the LDH-A cDNA clones from human and mouse (Tsujibo *et al.*, 1985; Akai *et al.*, 1985). In the present paper we present the exon–intron organization of human LDH-A functional gene and the genomic complexity of LDH-A gene-related sequences.

MATERIALS AND METHODS

The human genomic library was kindly provided by Dr. T. Maniatis (Lawn *et al.*, 1978) and the genomic clones containing LDH-A gene-related DNA sequences were isolated by plaque hybridization (Benson & Davis, 1977) using a ^{32}P-labelled probe of mouse LDH-A cDNA (Rigby *et al.*, 1977; Akai *et al.*, 1985). The DNA purified from the positive genomic clones as well as placenta was further analysed by restriction endonuclease cleavage and Southern blotting with the LDH-A cDNA probes from human and mouse (Southern, 1975; Maniatis *et al.*, 1982; Tsujibo *et al.*, 1985; Akai *et al.*, 1985). The isolated DNA fragments were further cleaved and subcloned into M13 mp10 or mp11 phages (Messing *et al.*, 1981). The M13 phages exhibiting positive hybridization to human LDH-A cDNA probe were isolated and nucleotide

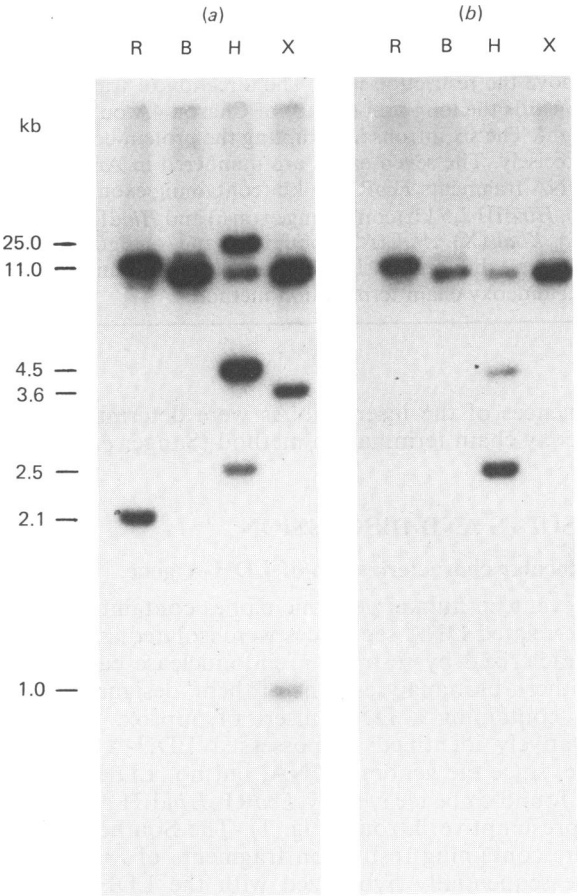

Fig. 1. Southern-blot analysis of human genomic clone λH448

The DNA fragments cleaved by restriction endonucleases *Eco*RI (R), *Bam*HI (B), *Hin*dIII (H) or *Xba*I (X) were electrophoresed on 0.75% agarose gel and Southern-transferred to nitrocellulose. The blotting filters were sequentially hybridized with the LDH-A cDNA probes of protein-coding sequence from human (*a*) and mouse (*b*). The estimated sizes (kb) of the DNA fragments hybridized to the LDH-A protein-coding probes are given. The DNA fragments containing the 3′ untranslated region exhibited positive hybridization with human LDH-A cDNA probe, but not with mouse probe because of sequence divergence (results not shown).

Abbreviations used: LDH, lactate dehydrogenase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; kb, kilobases; bp, base pairs.

* To whom correspondence and reprint requests should be addressed.

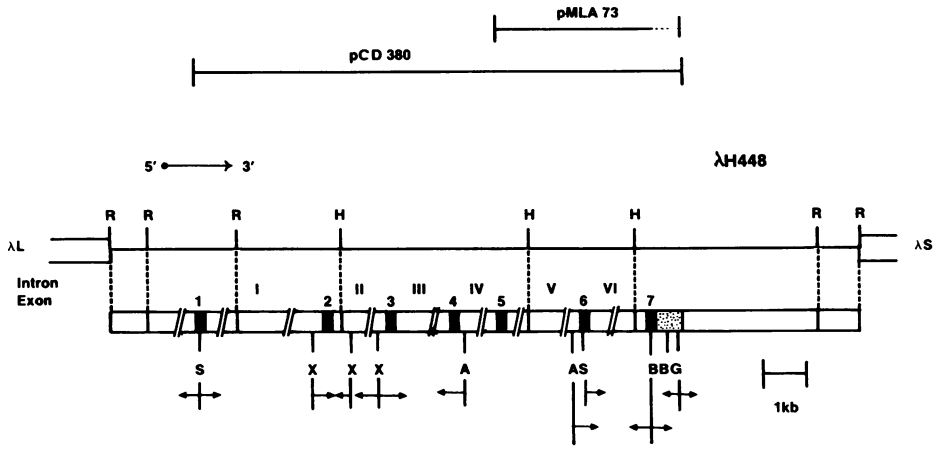

Fig. 2. Restriction endonuclease map and nucleotide sequencing strategy of human LDH-A gene

The restriction endonuclease map of human genomic clone λH448 was deduced from the Southern-blotting results presented in Fig. 1 as well as those of double cleavage. The LDH-A cDNA probes of human pCD380 and mouse pMLA73 are indicated above the restriction map. The direction of transcription is denoted by an arrow. The coding region of the LDH-A oriented towards the long arm (20 kb) of Charon 4A on the left, and the 3′ untranslated region towards the short arm (10.5 kb) on the right. The six introns interrupting the protein-coding sequence are indicated by Roman numerals and their sizes are not known precisely. The seven exons are numbered in Arabic and shown by solid blocks with the 3′ untranslated region stippled. The DNA fragments *Eco*RI 2.1 kb (containing exon 1), *Xba*I 1.0 kb (containing exon 2), *Hin*dIII 4.5 kb (containing exons 3, 4 and 5), *Hin*dIII 2.5 kb (containing exon 6) and *Hin*dIII–*Eco*RI 4.3 kb (containing exon 7) were isolated and further cleaved by *Sau*3AI (S), *Xba*I (X), *Alu*I (A), *Bam*HI (B) and/or *Bgl*II (G), and cloned in M13 mp10 or mp11. The DNAs were purified from M13 phages exhibiting positive hybridization to human LDH-A cDNA probe and their nucleotide sequences were determined by the dideoxy chain termination method.

sequences of the insert DNAs were determined by the dideoxy chain termination method (Sanger *et al.*, 1977).

RESULTS AND DISCUSSION

Molecular characterization of LDH-A gene

Seventeen human genomic clones containing LDH-A gene-related DNA sequences were isolated and partially characterized by restriction endonuclease cleavage and Southern-blot analyses. One of them, designated λH448 and containing a DNA insert of approx. 17 kb, was tentatively identified to possess a LDH-A functional gene, since the genomic DNA, but not LDH-A cDNA, was found to be cleaved by *Eco*RI, *Hin*dIII and *Xba*I at its presumptive introns (Fig. 1). The Southern-blotting filters containing restriction fragments of λH448 DNA were sequentially hybridized with the LDH-A cDNA probes from mouse and human (Akai *et al.*, 1985; Tsujibo *et al.*, 1985). The mouse pMLA73 LDH-A cDNA probe localized the three exons of the *C*-terminus because it contains only the portion of mRNA coding for amino acids 201–331 and the 3′ untranslated region. The human pCD380 LDH-A cDNA probe, which contains the entire protein-coding sequence of 999 bp as well as the 5′ (97 bp) and 3′ (565 bp) untranslated regions, identified the additional four exons of the *N*-terminus. The orientation of the LDH-A gene in the λH448 clone, the restriction endonuclease map and the DNA sequencing strategy are indicated in Fig. 2. The partial nucleotide sequences of λH448 genomic DNA were aligned with that of human pCD380 LDH-A cDNA (Tsujibo *et al.*, 1985). The protein-coding sequence of the human LDH-A gene is interrupted by six introns at codon numbers 41–42, 81, 139, 197, 236 and 277–278 (Fig. 3). The determined exon–intron junctions appear to follow the consensus splicing sequence of G-T at 5′ donor site and A-G at 3′ acceptor site (Breathnach *et al.*, 1978; Sharp, 1981). The nucleotide sequence of exons determined from genomic clone λH448 was found to be identical with that of LDH-A cDNA clone pCD380 (Tsujibo *et al.*, 1985) except that the published sequence of two nucleotides present in the 3′ untranslated region of LDH-A cDNA should be corrected as underlined. The seven protein-coding exons have similar sizes (150, 118, 174, 176, 118, 124 and 162 bp), while the sizes of the six introns were not known precisely. These six introns of human LDH-A gene are located at the same positions as those of mouse LDH-A gene (Li *et al.*, 1985). All six introns are positioned at random-coil regions or near the ends of secondary structures on the surface of monomeric subunit, although those for the third and fourth introns are near proline residues at internal positions in the tetrameric LDH molecule.

The sequence of 24 nucleotides 5′ to the presumptive ATG translation initiation site, but not the sequence of −97 to −25, of human pCD380 LDH-A cDNA is identical with that of genomic clone λH448, indicating the presence of an additional intron in the 5′ untranslated region of LDH-A gene. The exact transcription initiation site and its 5′ upstream control sequence remain to be determined experimentally. The nucleotide sequence of the 3′ untranslated region obtained from genomic clone λH448 concurs with that of human LDH-A cDNA sequence. This result indicates that the 3′ untranslated region is not interrupted by an intron and the poly(A) addition site is located 13 nucleotides 3′ to the AATAAA signal (Proudfoot & Brownlee, 1976). In short, human

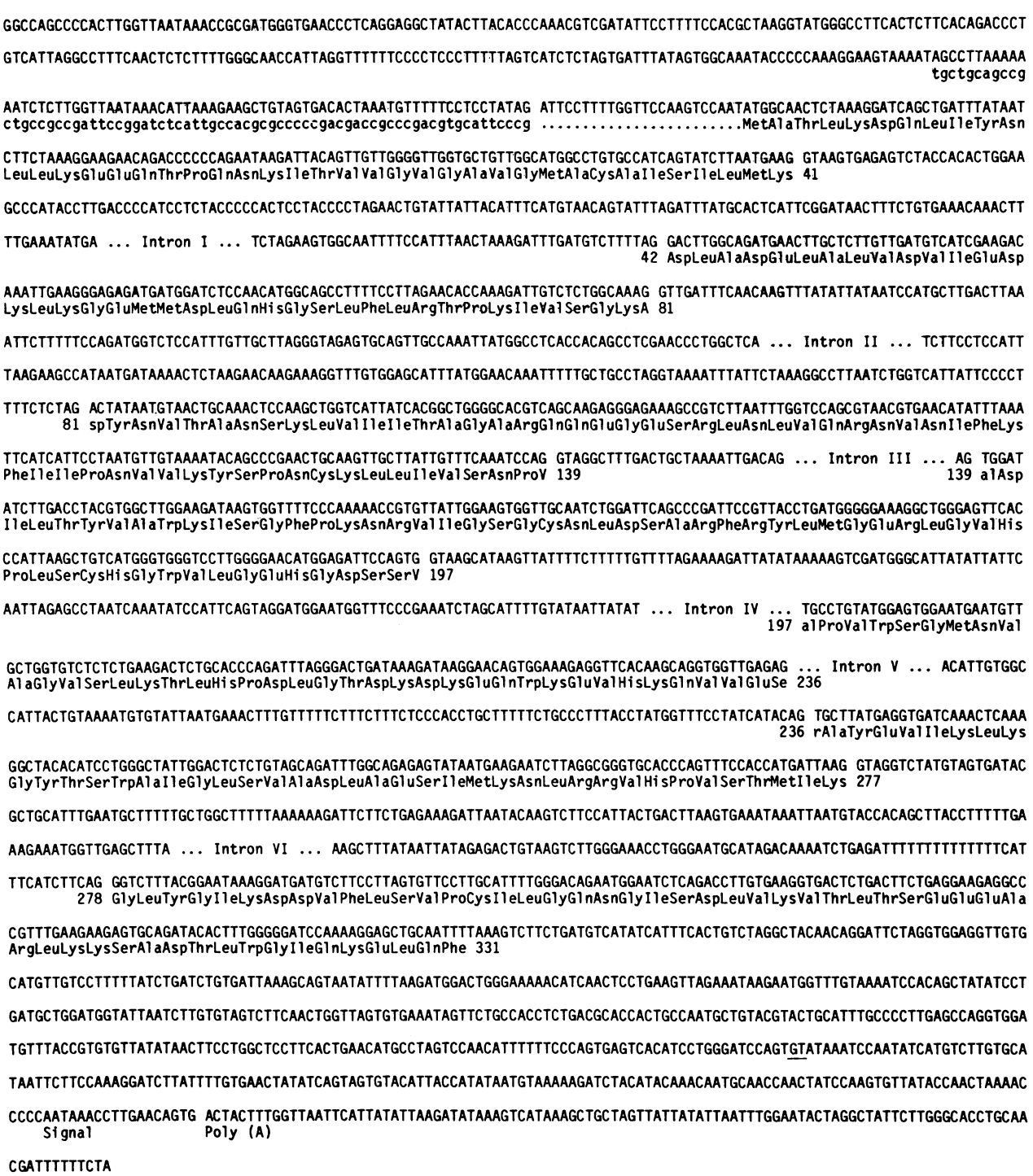

Fig. 3. Nucleotide sequences of exons and their flanking regions of the human LDH-A gene

The sequence of 24 nucleotides 5′ to the translation initiation ATG of LDH-A cDNA is identical with that of the exon 1 from genomic clone λH448 and is indicated by dots. The nucleotide sequence of −97 to −25 of LDH-A cDNA has not yet matched with that of genomic clone λH448 and is shown in lower case. Although the exon 5 (codons 197–236) must be present in the *Hin*dIII fragment of 4.5 kb because of positive hybridization with mouse LDH-A cDNA coding for amino acids 201–331, its nucleotide sequence was not determined experimentally. The exon–intron junctions of exon 5 were deduced from those of exons 4 and 6, and the presumptive sequence of exon 5 is that of human LDH-A cDNA clone pCD380.

LDH-A functional gene spans approx. 12 kb beginning at the translation initiation ATG and extending to the poly(A) addition site. The information on the LDH-A gene structure will facilitate the molecular characterization of genetic mutations affecting mammalian LDH-A_4 isoenzymes (Soares, 1979; Charles & Pretsch, 1981; Kanno *et al.*, 1983).

As discussed previously (Rossmann *et al.*, 1975; Li *et al.*, 1983*a*, 1985), all dehydrogenases consist of the coenzyme-binding domain and different catalytic do-

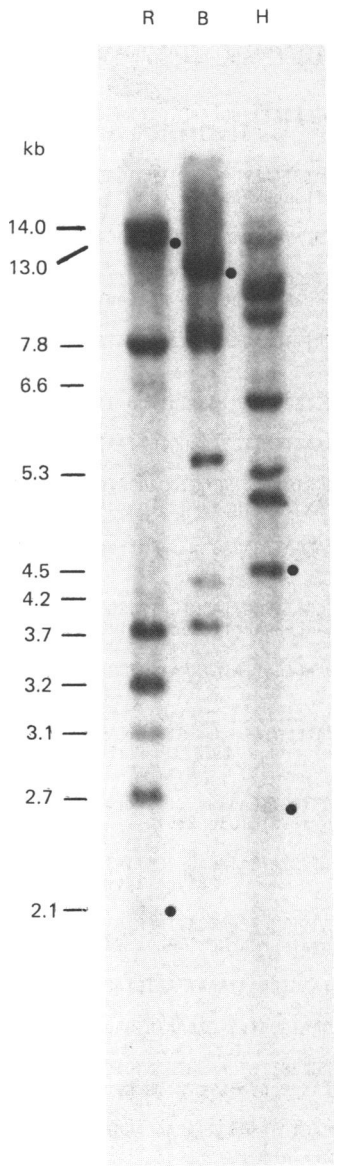

Fig. 4. Southern genomic blot analysis of human LDH-A gene-related sequences

The high-M_r DNA was isolated from human placenta and cleaved with restriction endonucleases *Eco*RI (R), *Bam*HI (B) or *Hin*dIII (H). After electrophoresis on 0.75% agarose gel, the DNA fragments were transferred to a nitrocellulose filter and hybridized with the human LDH-A cDNA probes. The result shown in the Figure was obtained from the hybridization with human LDH-A protein-coding probe in 6 × SSC at 63 °C overnight and washed twice with 3 × SSC at 42 °C for 30 min each. The estimated sizes (kb) of the *Eco*RI fragments exhibiting positive hybridization are given. Similar multiple hybridization bands were also observed with the 3′ untranslated sequence probe and the result is not shown.

mains, and the NAD-binding domain is composed of two similar mononucleotide binding folds. It is of interest to compare the exon–intron organization of the coenzyme-binding domains from mammalian LDH-A genes with that of chicken GAPDH gene (Stone *et al.*, 1985). An intron was found to separate the coenzyme-binding and catalytic domains of the GAPDH gene, but not the LDH-A gene. The adenine and nicotinamide binding folds of both LDH-A and GAPDH genes are divided by an intron at the corresponding positions. However, both mononucleotide-binding folds are further interrupted by introns at different positions of the LDH-A and GAPDH genes. While the entire 3′ untranslated regions of both human LDH-A and chicken GAPDH genes are not interrupted, their 5′ untranslated sequences are interrupted by an intron. The future investigation on the exon–intron organization of other dehydrogenase genes, including LDH-B and LDH-C genes, will certainly yield interesting insight on the molecular evolution and gene regulation of these genes.

Genomic complexity of LDH-A gene-related sequences

The Southern-blot analysis of human placenta DNA with human LDH-A cDNA probes demonstrates the presence of multiple LDH-A gene-related sequences (Fig. 4). The Mendelian inheritance of genetic mutations resulting in the LDH-A deficiency indicates only a single functional LDH-A gene expressed in human genomes (Kanno *et al.*, 1983). The DNA fragments *Eco*RI 14 kb, *Eco*RI 2.1 kb, *Bam*HI 11 kb, *Hin*dIII 4.3 kb and *Hin*dIII 2.5 kb correspond to those of the human LDH-A functional gene present in λH448 clone. As shown previously (Tsujibo *et al.*, 1985), the *Eco*RI 3.2 kb fragment found in seven overlapping genomic clones contains a processed LDH-A pseudogene. The partial characterization of the *Eco*RI fragments of 3.7 kb and 3.1 kb isolated from genomic clones indicates that both DNA fragments contain different processed LDH-A pseudogenes (S.S.-L. Li, unpublished work). Thus, the *Eco*RI fragments of 3.7, 3.2 and 3.1 kb observed in Southern-blot analysis of placenta DNA are obtained from three different processed LDH-A pseudogenes. The *Eco*RI fragments of 13, 7.8, 6.6, 5.3, 4.5 4.2 and 2.7 kb may contain either other LDH-A pseudogenes or LDH-B/LDH-C gene-related sequences. The human LDH-A functional gene has been genetically mapped on the short arm of chromosome 11. However, the chromosomal locations of these LDH-A pseudogenes remain to be determined. It has been proposed that the intronless pseudogenes were derived from the processed mRNA intermediates through retrovirus-like transposable elements (Lueders *et al.*, 1982), since the pseudogenes are often flanked by repeated DNA sequences. It may be also noted that a much higher multiplicity of GAPDH gene-related sequences has been shown in mammalian genomes (Arcari *et al.*, 1984; Benham *et al.*, 1984; Hanauer & Mandel, 1984; Piecharczyk *et al.*, 1984; Tso *et al.*, 1985).

We thank Mr. H. F. Tiano for technical assistance, Dr. T. Maniatis for the human genomic library, Dr. B. H. Judd for his interest and support to this project, and Drs. C. Aquadro and L. Skow for reading the manuscript.

REFERENCES

Akai, K., Yagi, H., Tiano, H. F., Pan, Y.-C. E., Shimizu, M., Fong, K., Jungmann, R. A. & Li, S. S.-L. (1985) Int. J. Biochem., **17**, 645–648

Arcari, P., Martinelli, R. & Salvatore, F. (1984) Nucleic Acids Res. **12**, 9179–9189

Benham, F. J., Hodgkinson, S. & Davies, K. E. (1984) EMBO J. **3**, 2635–2640
Benson, W. D. & Davis, R. W. (1977) Science **196**, 180–182
Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) Proc. Natl. Acad. Sci. U.S.A. **75**, 4853–4857
Charles, D. J. & Pretsch, W. (1981) Biochem. Genet. **19**, 301–309
Hanauer, A. & Mandel, J. L. (1984) EMBO J. **3**, 2627–2633
Kanno, T., Sudo, K., Kitamura, M., Miwa, S., Ichiyama, A. & Nishmura, Y. (1983) Curr. Top. Biol. Med. Res. **7**, 131–150
Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) Cell **15**, 1157–1174
Li, S. S.-L., Fitch, W. M., Pan, Y.-C. E. & Sharief, F. S. (1983*a*) J. Biol. Chem. **258**, 7029–7032
Li, S. S.-L., Feldmann, R. J., Okabe, M. & Pan, Y.-C. E. (1983*b*) J. Biol. Chem. **258**, 7005–7016
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
Lueders, L., Leder, A., Leder, P. & Kuff, E. (1982) Nature (London) **295**, 426–428
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Markert, C. L., Shakelee, J. B. & Whitt, G. S. (1975) Science **189**, 102–114
Matrisian, L. M., Rautmann, G., Magun, B. E. & Breathnach, R. (1985) Nucleic Acids Res. **13**, 711–726
Messing, J., Crea, R. & Seeberg, P. H. (1981) Nucleic Acids Res. **9**, 309–321
Miles, M. F., Hung, P. & Jungmann, R. A. (1981) J. Biol. Chem. **256**, 12545–12552
Pan, Y.-C. E., Sharief, F. S., Okabe, M., Huang, S. & Li, S. S.-L. (1983) J. Biol. Chem. **258**, 7005–7016
Piecharczyk, M., Blanchard, J. M., Riaad-El, S., Dani, C., Marty, L. & Jeanteur, P. (1984) Nature (London) **312**, 469–471
Proudfoot, N. J. & Brownlee, G. (1976) Nature (London) **263**, 211–214
Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. **113**, 237–251
Rossmann, M. G., Liljas, A., Branden, C. L. & Banaszak, L. J. (1975) Enzymes 3rd Ed. **11**, 61–102
Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Sharp, P. A. (1981) Cell **23**, 643–646
Soares, E. R. (1979) Environ. Mutagen. **1**, 19–25
Southern, E. M. (1975) J. Mol. Biol. **98**, 503–517
Stone, E. M., Rothblum, K. N., Alevy, M. C., Kuo, T. M. & Schwartz, R. J. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1628–1632
Tso, J. Y., Sun, X.-H., Kao, T-h., Reece, K. S. & Wu, R. (1985) Nucleic Acids Res. **13**, 2485–2502
Tsujibo, H., Tiano, H. F. & Li, S. S.-L. (1985) Eur. J. Biochem. **147**, 9–15
Williams, K. R., LoPresti, M. B., Stone, K. L. & Patel, G. (1984) Fed. Proc. Fed. Am. Soc. Exp. Biol. **43**, 1541 (abstr.)

Received 4 July 1985; accepted 22 July 1985