Isolation and characterization of genomic clones corresponding to the human type II procollagen gene

Charles M. Strom* and William B. Upholt⁺

*+Departments of Pediatrics and +Biochemistry, Joseph P. Kennedy, Jr. Mental Retardation Research Center, University of Chicago, Pritzker School of Medicine, Hospital Box 413, 5841 South Maryland Avenue, Chicago, IL 60637, USA

Received 17 October 1983; Revised and Accepted 6 December 1983

ABSTRACT

A recombinant human DNA library was screened using probes corresponding to the chick $\alpha 1(\text{II})$ procollagen gene. This resulted in the isolation of 2 different genomic clones, LgHCol(II)a and LgHCol(II)b. LgHCol(II)a was identified as corresponding to the $\alpha 1(\text{II})$ gene by comparative hybridization and DNA sequence analysis. DNA sequence established that LgHCol(II)a extends at least from amino acid 694 of the triple helix through 54 amino acids of the COOH-propeptide. Hybridization with a probe containing only the exon at the 3' end of the chicken gene suggests that the clone contains the 3' end of the human gene. Thus LgHCol(II)a contains approximately 40% of the coding sequences of the human type II collagen gene.

INTRODUCTION

Type II collagen is the predominant collagen of cartilage. The molecule is secreted by chondrocytes into the extracellular space, associating with other cartilage components, to stabilize endochondral skeletal structures (1,2).

At least 5 collagen types have been identified, comprised of at least 9 separate collagen gene products (2). Type II collagen is comprised of 3 $\alpha 1(\text{II})$ chains. These are synthesized as larger procollagen chains which contain N-terminal and C-terminal amino acid sequences called propeptides. These are thought to be involved in secretion, fibrillogenesis, or alignment of the three pro $\alpha 1(\text{II})$ chains during triple helix formation (2,3). After secretion into the extracellular matrix the propeptides are cleaved, forming the mature type II collagen molecule (2,3).

Upholt and co-workers have successfully isolated cDNA clones and overlapping genomic clones corresponding to the

chick α 1(II) procollagen gene (4,5). We used recombinant DNA probes corresponding to the COOH-propeptide region of the chick α 1(II) procollagen gene to screen a bacteriophage lambda recombinant library of human DNA. This resulted in the isolation of 2 different lambda clones corresponding to the human α 1(II) procollagen gene. LgHCol(II)a has been directly identified by comparative hybridization and DNA sequencing studies. DNA sequence analysis shows that this clone contains sequences coding for the region from amino acid 694 of the triple helical region through 54 amino acids of the COOH-propeptide. Hybridization studies using chicken probes suggest that the 3' end of the gene is contained within LgHCol(II)a.

MATERIALS AND METHODS

Restriction endonucleases and other enzymes were purchased from Bethesda Research Laboratories, New England Biolabs, and Promega Biotec. Radiochemicals were obtained from Amersham. Nitrocellulose and DEAE filters were obtained from Schleicher and Schuell.

Hybridization Probes

Purified inserts from the cDNA plasmids pCAR1 and pCAR2 which contain DNA corresponding to the 3' end of the chick type II procollagen mRNA (4) were used for screening the genomic library.

Two purified HinfI fragments from the subcloned 5.0kb BamHI-EcoRI fragment of the chicken genomic clone LgCOL(II) (5,6) were also used for the screening and mapping of the human genomic clone. A 450 bp HinfI fragment has a 289 bp exon coding for the last 15 amino acids of the triple helix, 27 amino acids of the COOH-telopeptide and 54 amino acids of the COOH-propeptide, was used in screening. This exon is the fourth exon from the 3' end of the chick α 1(II) procollagen gene and probe derived from this region will be referred to as exon 4 probe. A 400 bp HinfI fragment containing a 162 bp exon coding for amino acids 712-765 of the triple helical region of chicken type II collagen was used in mapping the extent of the human genomic clones.

The cDNA plasmids pCAL1 and pCAL2 which have sequences

corresponding to the $\alpha 1(I)$ procollagen and $\alpha 2(I)$ procollagen mRNAs respectively were also used in hybridization experiments (4,7). Plasmid pCAL1 contains sequences coding for the last 191 amino acids of the chicken $\alpha 1(I)$ COOH-propeptide and pCAL2, which corresponds exactly to pCOL1 of Sobel *et al.* (8) contains sequences coding for 63 amino acids near the COOH end of the $\alpha 2(I)$ COOH-propeptide. Inserts and plasmids were labeled by nick translation (9) in the presence of [α - ^{32}P] dTTP, dCTP, and dGTP.

Screening of Human Genomic Library

The human AluI/HaeIII partial digestion lambda Charon 4A library was generously donated by Dr. Tom Maniatis. The construction and description of the library is described elsewhere by Fritsch *et al.* (10). Purified pCAR1 insert and the exon 4 probe were used to screen the library, using the procedure of Benton and Davis (11).

Restriction Endonuclease Digestions

Restriction digestions were carried out under conditions recommended by the manufacturers.

Hybridization Conditions

Hybridizations for library screening were performed at 67° for 14-17 hours in 6XSSC (1XSSC is 0.15M NaCl, 15mM sodium citrate), 0.1% SDS and 10X Denhardt's solution (12). Southern filter hybridizations (13) were carried out at various temperatures for 14-20 hours in 50% formamide, 6XSSC, 0.01 M EDTA, 5X Denhardt's solution, 0.5% SDS and 100 $\mu\text{g}/\text{ml}$ denatured salmon sperm DNA. The filters were washed at room temperature with 5 changes of 2xSSC, 0.1% SDS, followed by 4 changes of 1xSSC, 0.1% SDS at 5° below the hybridization temperature.

DNA Preparation and Characterization

Phage DNA was prepared by the plate lysate method of Fritsch, as described by Maniatis *et al.* (14). Plasmid DNA was prepared by the boiling method of Holmes and Quigley (14,15). Two fragments of LgHCol(II)a were subcloned into pBR322. LgHCol(II)a was sequentially digested with BamHI and EcoRI and a 2.0 kbp fragment identified by Southern hybridization as corresponding to exon 4 was recovered from a 1% horizontal agarose gel with DEAE paper using the procedure

described by the manufacturer (16). A sequential BamHI/EcoRI digestion of pBR322 was followed by Sepharose 4B column chromatography to separate the 375 bp fragment from the remainder of the plasmid (8). The large pBR322 fragment and the 2.0 kbp fragment were then co-precipitated and ligated using T4 DNA ligase. E.coli HB101 was transformed with the ligated DNA and ampicillin resistant clones were picked. The resulting recombinant clone was named pgHCol(II)A. Similarly, pgHCol(II)B, containing a 2.7 kbp EcoRI fragment corresponding to the 5' end of the insert in the genomic clone, was constructed by insertion of the purified EcoRI fragment into pBR322 which had been EcoRI digested and treated with bacterial alkaline phosphatase.

DNA Sequencing

DNA restriction fragments were isolated from polyacrylamide gels and labeled at 5' ends with [γ - 32 P]ATP, using T4 polynucleotide kinase or at recessed 3' ends of restriction endonuclease cleavage sites using the large fragment of E.coli DNA polymerase I and the appropriate [α - 32 P]deoxynucleotide triphosphate. Sequencing was performed as described by Maxam and Gilbert (17). All sequences described were determined on both strands.

RESULTS

Isolation of Human Type II Procollagen Genomic Clones

Approximately 1.5×10^6 plaques of the human genomic library were screened. Duplicate filters were hybridized with a mixture of two chicken type II procollagen recombinant DNA probes, pCAR1 insert and exon 4 genomic probe. Sixty positive plaques were obtained. Thirty were re-screened, resulting in 7 positively hybridizing plaque purified phage. In the plaque-purification step, triplicate filters were made from each plate. Two filters from each plate were hybridized with pCAR1 probe and the third filter was hybridized with the exon 4 probe. Filters from these plaques hybridized to both probes. No plaques were observed which hybridized with only one of the two probes.

Restriction Fragment Analysis of Genomic Clones

Small scale preparations of bacteriophage DNA were carried out by the plate lysate method of Fritsch (14). Digestions with EcoRI, BamHI and sequential digestions with EcoRI and BamHI revealed two different digestion patterns among the seven clones. Six of the clones demonstrated EcoRI bands at 20 kbp (lambda), 10.5 kbp (lambda), 4.0 kbp, 3.7 kbp, 2.7 kbp, and 0.87 kbp. These clones have been named LgHCol(II)a. The final clone differed only in the presence of a 4.7 kbp band and a 3.0 kbp in place of the 0.87 kbp band and the 2.7 kbp band (Fig. 1). This clone is named LgHCol(II)b. Subsequent mapping revealed that the 0.87 kbp fragment in the LgHCol(II)a is the most 3' fragment and the 2.7 kbp fragment is the most 5' fragment (with respect to mRNA polarity) adjoining the lambda vector. The estimated lengths of the genomic inserts in LgHCol(II)a and LgHCOL(II)b are therefore approximately 11.2 kbp and 15.4 kbp respectively. Southern blots of the EcoRI, BamHI, and sequential EcoRI and BamHI digestions hybridized with the exon 4 genomic probe revealed identical bands in LgHCol(II)b and all clones of LgHCol(II)a. The shown

Fig. 1. Ethidium bromide stained agarose gel of isolated clones. Lane 1, LgHCol(II)b; lanes 2-7, six isolated identical clones of LgHCol(II)a.

Fig. 2. Hybridization of chicken $\alpha 1(\text{II})$ exon 4 genomic probe to various restriction fragments of human genomic clones. A. Lanes 1-3, 4-6, 7-9, 10-12, and 13-15, five independent isolates of LgHCol(II)a; lanes 16-18, LgHCol(II)b. Lanes 1, 4, 7, 10, 13, and 16, EcoRI digests; lanes 2, 5, 8, 11, 14, and 17, EcoRI/BamHI digests; and lanes 3, 6, 9, 12, 15, and 18, BamHI digests. B. LgHCol(II)b, Lane 7, and LgHCol(II)a, lanes 1-6, digested with SmaI.

probe hybridized to a 3.7 kbp EcoRI fragment, a 4.2 kbp BamHI fragment and a 2.0 kbp BamHI/EcoRI fragment. (Fig. 2A). Similarly, hybridization of the exon 4 probe to Southern blots of HinfI and SmaI digestions resulted in single bands of 600 bp and 1.7 kbp respectively (Fig. 2B and Fig. 3A). All the above hybridizations were carried out at 45° in 50% formamide. Identification of Genomic Clones

The identity of LgHCol(II)a was confirmed by comparing the stability of hybrids formed with probes corresponding to chick $\alpha 1(\text{II})$, $\alpha 1(\text{I})$, and $\alpha 2(\text{I})$ procollagen genes. From work in the chick system, it is known that the $\alpha 1(\text{II})$ and $\alpha 1(\text{I})$ procollagen genes share greater sequence homology than any other pair of known collagen types (4,5). In analyzing chick genomic blots, the stringent hybridization conditions of 50% formamide and 51° are necessary to eliminate the cross hybridization between chicken $\alpha 1(\text{I})$ and $\alpha 1(\text{II})$ DNA sequences (4). In the following experiments, all hybridizations were carried out in 50% formamide. Initially, DNA from LgHCol(II)a was digested with HinfI and the resulting filters were hybridized with the exon 4 probe of the chick $\alpha 1(\text{II})$ gene at 45°, reveal-

Fig. 3. Thermal stability of various collagen gene probes hybridized with *Hinf*I fragments of LgHCol(II)a and LgHCol(II)b. A. Lanes 1-6, six clones of LgHCol(II)a; lane 7, LgHCol(II)b hybridized with chicken α (II) exon 4 probe. B. Same filter as in A hybridized with chicken α (I) probe pCAL1. C. LgHCol(II)a hybridized at 45° with a combination of exon 4 (α (II)) and pCAL1 (I(I)) probes. D. Duplicate filter as in C hybridized at 51° with the same combination of α (I) and α (II) probes.

ing a single band of 600 bp (Fig. 3A). When this filter was washed and re-hybridized at 45° with a chick α (I) probe (pCAL1) (4), a single band at 1600 bp was observed (Fig. 3B). This cross hybridization at 45° is expected due to the sequence homology between the α (I) and α (II) sequences. The two probes hybridized to fragments of different sizes because the α (I) and α (II) probes differ slightly in their locations in their respective genes. This difference was exploited in the following experiment to compare the stabilities of these hybrids. DNA from LgHCol(II)a was digested with *Hinf*I, electrophoresed on an agarose gel, and transferred to nitrocellulose. Duplicate filters were hybridized at various temperatures with a mixture of 4×10^6 cpm of α (II) probe made from the exon 4 fragment and 16×10^6 cpm of α (I) probe prepared from whole pCAL1 plasmid. At 43°, both the 1600 bp band (α (I) hybridization (Fig. 3C)) and the 600 bp band (α (II) hybridization) were observed. At 51°, however, only the 600 bp band corresponding to α (II) hybridization was observed (Fig. 3D). This experiment showed that LgHCol(II)a contains sequences more homologous to the

Fig. 4. Restriction enzyme cleavage maps of LgHCol(II)a and LgHCol(II)b. A, LgHCol(II)a; B, LgHCol(II)b; C, subclone pgHCol(II)B; and D, subclone pgHCol(II)A. Restriction enzymes are EcoRI, E, and BamHI, B. 5'-3' orientation with respect to the mRNA and λ vector arms are also indicated.

chicken α 1(II) probe than to the α 1(I) probe. No hybridization at 45° was observed using the chick α 2(I) cDNA probe pCAL2 (data not shown). Since the α 1(I) and α 1(II) genes show the greatest sequence similarity of the collagen genes, these experiments indicate that the isolated clones most likely represent the α 1(II) procollagen gene.

The amino acid sequences of the COOH-telopeptides and COOH-propeptides of collagen molecules provide an excellent means of identification of collagen types. Sequence analysis from recombinant DNA studies shows that exon 4 is the least conserved of the exons at the 3' end of collagen genes (6,18-20). Therefore, analysis of the DNA sequence of this exon should yield reasonable identification of the clones. For this purpose, a 2.0 kbp BamHI/EcoRI fragment was subcloned into pBR322 as described in MATERIALS AND METHODS. The location of subclone pgHCol(II)A is shown in Fig. 4. The DNA sequence was determined beginning at the BamHI and extending 165 bases upstream (Fig. 5). The sequence in Fig. 5 contains 154 bases which code for 51-1/3 amino acids of exon 4 of the human type II procollagen gene and 11 bases of intron. The split codon at the 3' end of the exon also occurs in the chicken α 2(I), α 1(II) and α 1(III) procollagen genes (6,20,21). Since the amino acid sequence of human type II

		30 ^C															40 ^C				
pgHCol(II)A	na	GCC	GGT	GGC	CTG	AGA	CAG	CAT	GAC	GCC	GAG	GTG	GAT	GCC	ACA						
	aa	ALA	GLY	GLY	LEU	ARG	GLN	HIS	ASP	ALA	GLU	VAL	ASP	ALA	THR						
chick α1(II)	na		A	G		C	G		C					C							
	aa									VAL											
chick α1(I)	na	AAC	TG	A	C	C	G	C	GC	CTG				C	A						
	aa	ASN	VAL	MET					ASP	ARG	LEU				THR						
		50 ^C																			
pgHCol(II)A	na	CTC	AAG	TCC	CTC	AAC	AAC	CAG	ATT	GAG	AGC	ATC	CGC	AGC	CCC						
	aa	LEU	LYS	SER	LEU	ASN	ASN	GLN	ILE	GLU	SER	ILE	ARG	SER	PRO						
chick α1(II)	na		A					T													
	aa																				
chick α1(I)	na	G		AG		G	C	A		C		A									
	aa					SER	GLN					ASN									
		60 ^C															70 ^C				
pgHCol(II)A	na	GAG	GGC	TCC	CGC	AAG	AAC	CCT	GCT	CGC	ACC	TGC	AGA	GAC	CTG						
	aa	GLU	GLY	SER	ARG	LYS	ASN	PRO	ALA	ARG	THR	CYS	ARG	ASP	LEU						
chick α1(II)	na				AAG					C	A	G		C							
	aa				LYS																
chick α1(I)	na	A		A					G	C			C	G							
	aa			THR																	
		79 ^C																			
pgHCol(II)A	na	AAA	CTC	TGC	CAC	CCT	GAG	TGG	AAG	AGT	G	gtaagcttggg									
	aa	LYS	LEU	CYS	HIS	PRO	GLU	TRP	LYS	SER											
chick α1(II)	na				T	C					C										
	aa																				
chick α1(I)	na	G	A	G		GGC	C				C										
	aa			MET		GLY	ASP														

Fig. 5. DNA sequence of exon 4 region in LgHCol(II)a DNA. Sequence was determined extending 5' from the BamHI site in pgHCol(II)A.

This sequence and the derived amino acid sequence are compared with the corresponding regions of the chicken α1(II) (6) and α1(I) genes (17). The numbering corresponds to the chicken α1(I) sequence numbering sequentially from the first amino acid of the COOH-telopeptide.

procollagen is not known, the identification was made by comparisons with known chick DNA sequence (6,18). The coding sequence of pgHCol(II)A differs in only 3 of 51 amino acids when compared with the chick type II procollagen sequence for the corresponding region, whereas the sequence differs in 14 of 51 amino acids when compared with the chick α1(I) sequence. An analogous comparison between human and chick α2(I) collagen sequence for this same region reveals 2 substitutions out of 51 amino acids (19) (see TABLE 1). These results show clearly that the isolated clones correspond to human α1(II) sequences.

Determination of the Extent of LgHCol(II)a

The BamHI-EcoRI restriction map of the entire length of LgHCol(II)a was generated by restriction mapping and by hy-

TABLE 1

Comparison of Nucleotide and Amino Acid Sequence Similarities
for Various Collagen Types in the Exon 4 Region

	nucleotides	amino acids
pgHCol(II)A: chick α 1(type II)	130/154=84%	48/51=94%
pgHCol(II)A: chick α 1(type I)	114/154=74%	37/51=72%
chick α 1(I): chick α 1(II)	117/154=76%	35/51=69%
human α 2(I): chick α 2(I)	134/154=87%	49/51=96%

Chicken α 1(I) and α 2(I) sequences are from Fuller and Boedtker (18), human α 2(I) sequence is from Bernard *et al.* (19), and chicken α 1(II) sequence is from Upholt *et al.* (6).

bridization of various chicken type II procollagen gene probes to Southern blots (Fig. 4). Hybridization of pCAR1 (a probe corresponding to 129 bp of exon 3, all of exon 2 and 144 bp of exon 1) to blots of EcoRI/BamHI digestions of LgHCol(II)a is shown in Fig. 2A. A 400 bp probe containing a 162 bp exon coding for amino acids 712-765 of the triple helical region of chick procollagen hybridized to a 4.0 kbp BamHI fragment (data not shown), a major EcoRI fragment of 4.0 kbp, a minor EcoRI fragment of 2.7 kbp, and two BamHI/EcoRI fragments of 2.2 and 1.8 kbp (Fig. 6A).

In order to elucidate the 3' extent of our clones in relationship to the type II procollagen gene, hybridizations were carried out using pCAR2, which contains only the first exon of the chick type II procollagen gene. This probe hybridizes to a 3.7 kbp EcoRI fragment and a 1.7 kbp EcoRI/BamHI fragment (Fig. 6B). This clearly indicates that the cross-hybridizing portion of exon 1 is contained in a 1.7 kbp internal BamHI/EcoRI fragment (see Fig. 4). LgHCol(II)a contains an additional 870 bp and LgHCol(II)b and an additional 4.7 kbp of sequence 3' to this fragment. If the human gene contains a 3' nontranslated sequence of similar length to the chick gene (approximately 500 bp), these clones should contain the 3' end of the human type II procollagen gene.

In order to establish the 5' extent of the clones, the sub-clone of the most 5' EcoRI fragment (pgHCol(II)B) was con-

Fig. 6. Determination of the 5' and 3' extents of LgHCol(II)a and LgHCol(II)b by hybridization with chicken probes. A. LgHCol(II)a (lanes 1 and 2) and LgHCol(II)b (lanes 3 and 4) digested with *Eco*RI (lanes 1 and 3) or *Eco*RI/*Bam*HI (lanes 2 and 4) and hybridized to a chick cDNA containing only exon 1. B. LgHCol(II)a (lanes 1 and 2) and LgHCol(II)b (lanes 3 and 4) were digested with *Eco*RI (lanes 1 and 3) or *Eco*RI/*Bam*HI (lanes 2 and 4), electrophoresed on an agarose gel, transferred and hybridized to a chicken genomic probe containing a 162 bp exon coding for amino acid 712-765.

structured as described in MATERIALS AND METHODS (Fig. 4). One exon sequence beginning near the *Bam*HI site of pgHCol(II)B is shown in Fig. 7. Comparison with the known chick α 1(I) amino acid sequence (22) identified this as corresponding to amino acids 694-711 of the triple helical region. This sequence contains an entire 54bp exon and both splice junction sequences. Therefore, LgHCol(II)a contains the 3' end of the type II procollagen gene and extends in a 5' direction at least as far as amino acid 694 of the triple helix.

pgHCol(II)B	gagggcttgaggttctcaccocgtctcctctccccacacag
	694 700
pgHCol(II)B	GGA GCC ACT GGA TTC CCT GGA GCT GCT GGT CGC GTT
a.a	GLY ALA THR GLY PHE PRO GLY ALA ALA GLY ARG VAL
Chick α 1(I)a.a	
	710
pgHCol(II)B	GGG CCC CCA GGT CCA GAG gtcaccc
a.a	GLY PRO PRO GLY PRO GLU
Chick α 1(I) a.a	SER

Fig. 7. DNA sequence of an exon in pgHCol(II)B coding for amino acids 694-711 of the triple helical region. DNA sequence was determined extending 3' from the *Bam*HI site in pgHCol(II)B. Comparison is made with known amino acid sequence of chick α 1(I) (19).

LgHCol(II)a thus contains approximately 40% of the coding sequences of the human type II procollagen gene.

DISCUSSION

This paper describes the isolation of a portion of the human type II procollagen gene. The identity of one of the two resulting clones has been established by both comparative hybridization and sequence analysis. The derived amino acid sequence for this gene represents the first of such data for this important protein. Preliminary results have revealed that the subclone pgHCol(II)A hybridizes to a single band on Southern blots of BamHI and HindIII digestions of human placental DNA, suggesting that there is a single copy of this gene in a haploid human genome. In addition, we have sequenced one exon and a portion of a second exon, as well as parts of the 3 adjoining introns, and have found no evidence of frame shift mutations or termination codons in coding regions or altered splice junction sequences. Taken together, these observations suggest that the cloned sequences are unlikely to represent a pseudogene. The region of the human type II procollagen gene coding for amino acid 694 of the triple helical region through exon 4 occupies 7.6 kbp of the human genome. The corresponding regions of the chicken α 1(II) and α 2(I) procollagen genes occupy approximately 2.9 kbp (Upholt, unpublished results) and 6.0 kbp (23) respectively. These results suggest that the small size of the chicken type II procollagen gene compared with the chicken α 2(I) procollagen gene does not reflect a characteristic specific for type II procollagen genes. However, these results may not prove to be representative of the entire human gene as only 20% of the gene is being compared in this case and hybridization of the chicken probe for an exon coding for amino acids 712-765 (see above) suggests that a single intron of approximately 2 kbp is located between this exon and the most 5' exon (amino acids 694-711) identified in LgHCol(II)a.

ACKNOWLEDGEMENTS

We gratefully acknowledge the expert technical assistance of Carol Chien, James Mensch, and Gloria Enriquez. This work

was supported by the Children's Research Grant 6-93536, the Sprague Memorial Grant 6-93113, the Institutional Grant of the American Cancer Society 6-95604, and Public Health Service Grants HD-04583 and HD-09402.

REFERENCES

1. Von der Mark, K. (1980) *Current Topics in Developmental Biology* 14, 199-225.
2. Bornstein, P. and Sage, H. (1980) *Ann. Rev. Biochem.* 49, 957-1003.
3. Olsen, B.R. (1981) *in: Cell Biology of Extracellular Matrix*, E.D. Hay, ed., 139-177 Plenum Press, NY.
4. Vuorio, E., Sandell, L., Kravis, D., Sheffield, V.C., Vuorio, T., Dorfman, A. and Upholt, W.B. (1982) *Nucleic Acids Res.* 10, 1175-1192.
5. Sandell, L.J., Yamada, Y., Dorfman, A., and Upholt, W.B. (1983) *J. Biol. Chem.* 258, 11617-11621.
6. Sandell, L.J., Prentice, H., and Kravis, D. and Upholt, W.B. (submitted).
7. Vuorio, E.I., Schaefer, I.M., Vuorio, T.K., Dorfman, A., and Upholt, W.B. (1983) *Acta Chem. Scand. B* (in press).
8. Sobel, M.E., Yamamoto, T., Adams, S.L., DiLauro, R., Avvedimento, V.E., de Crombrughe, B., and Pastan, I. (1978) *Proc. Natl. Acad. Sci. USA* 75, 5846-5850.
9. Maniatis, T., Jeffrey, A., and Kleid, D.G. (1975) *Proc. Natl. Acad. Sci. USA* 72, 1184-1185.
10. Fritsch, E.F., Lawn, R.M., and Maniatis, T. (1980) *Cell* 19, 959-972.
11. Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180-182.
12. Denhardt, D.T. (1966) *Biochem. Biophys. Res. Commun.* 23, 641-646.
13. Southern, E.M. (1979) *Methods Enzymol.* 68, 152-176.
14. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor.
15. Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.* 114, 193-200.
16. Winberg, G. and Marrarskjold, M.-L. (1980) *Nucleic Acids Res.* 8, 253-264.
17. Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
18. Fuller, F. and Boedtker, H. (1981) *Biochemistry* 20, 996-1006.
19. Bernard, M.P., Myers, J.C., Chu, M.-L., Ramirez, F., Eikenberry, E.F., and Prockop, D.J. (1983) *Biochemistry* 22, 1139-1145.
20. Yamada, Y., Kuhn, K., and de Crombrughe, B. (1983) *Nucleic Acids Res.* 11, 2733-2744.
21. Dickson, L.A., Ninomiya, Y., Bernard, M.P., Pesciotto, D.M., Parsons, J., Green, G., Eikenberry, E.F., de Crombrughe, B., Vogeli, G., Pastan, I., Fietzek, P.P., and Olsen, B.R. (1981) *J. Biol. Chem.* 256, 8407-8415.

22. Highberger, J.H., Corbett, C., Dixit, S.N., Yu, W., Seyer, J.M., Kang, A.H., and Gross, J. (1982) *Biochemistry* 21, 2048-2055.
23. Tate, V., Finer, M., Boedtker, H., and Doty, P. (1983) *Cold Spring Harbor Symposium on Quantitative Biology*, Vol. XLVII, 1039-1049.