Vol. 90, No. 3, 1979 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 12, 1979 Pages 832-841

NUCLEOTIDE SEQUENCE OF A MOUSE KAPPA LIGHT CHAIN cDNA CLONED IN A BACTERIAL PLASMID

K. Fong, A. Lui, and W. Salser

Department of Biology and Molecular Biology Institute
University of California, Los Angeles, California 90024

Received August 20,1979

SUMMARY

A mouse MOPC21 cDNA previously cloned in plasmid pMB9(Higuchi et al., Proc. Natl. Acad. Sci.73(1976)2136-2140; Wall et al., Nucleic Acid Res. 5(1978)3113-3128) and is designated pL21-3 has been extensively characterized. Cleavage of pL21-3 with HpaII has shown the insert to be 910 basepairs long, consistent with the length of the entire variable and constant regions and the untranslated regions. Digestion of pL21-3 with various restriction endonucleases has established that the insert sequence starts from parts of the 5' leader region and extends downstream to include the untranslated 3' terminus. 131 nucleotides in the variable region corresponding to amino acids 49-91 have been determined.

INTRODUCTION

General procedures have been established for the construction of recombinant clones from polyA-containing mRNA (1-4). Starting from purified mouse MOPC21 immunoglobin kappa light chain polyadenylated mRNA (5), a double-stranded cDNA copy was made and joined to a plasmid vector pMB9 using the poly dA-T tailing method (1). After transformat-on by these recombinant molecules, tetracycline-resistant E. coli colonies containing the kappa light chain mRNA sequences were identified by colony hybridization as described by Grunstein and Hogness (6). Restriction endonuclease mapping (7) and hybridization studies (5) have shown pL21-5 to be derived primarily from the 3' constant coding region and that pL21-4, with the smallest insert sequence of about 500 base pairs (bp), may represent sequence at the 5' end and possibly the terminal cap. pL21-1 through pL21-3, on the other hand, contain a much larger insert sequence and are found to include both the variable and constant coding regions in addition to the flanking non-coding sequences (5). Recently, sequence analysis has established 77 nucleotides at the 5' terminus of the variable coding region in pL21-1 (8)

0006-291X/79/190832-10$01.00/0

and the entire constant and 3' untranslated region of MOPC21 kappa light chain mRNA by a separate group (9). In the following we report that the insert in pL21-3 bears a restriction map similar to that of pL21-1 and that 131 nucleotides in the 5' variable coding region have been determined.

MATERIALS AND METHODS

Enzymes and Chemicals Restriction endonucleases were purchased from Bethesda Research Laboratory. Bacterial alkaline phosphatase was from Worthington Biochemical. T4 polynucleotide kinase was from PL-Biochemicals. Gamma ^{32}P-ATP at a specific activity of about 1 mCi/mmol was prepared by the authors according to the protocol of Maxam and Gilbert (10). Acrylamide and pure urea were from Eastman Kodak Company and Schwarz/Mann respectively. Chemicals used for partial degradation reactions in DNA sequence determination were dimethyl sulfate (Aldrich Chemical), hydrazine(Eastman) and piperidine(Mallinckrodt Chemical).

Isolation of DNA Fragments for Sequencing pL21-3 plasmid DNA was cleaved with restriction enzymes and the appropriate fragments were then isolated from a 6% or 8% polyacrylamide gel by the elution method as described by Maxam and Gilbert (10). A 60-70% recovery was generally obtained for DNA fragments ranging from 500-1,000 bp.

Dephosphorylation and ^{32}P-labeling of 5' Termini Bacterial alkaline phosphatase was used to remove the 5' terminal phosphate of DNA fragments in 0.01M Tris-HCL (pH 9.0) at 37C for 30 min. After extraction by liquified phenol and ether, the dephosphorylated DNA fragments were labeled at 5' ends with gamma ^{32}P-ATP using T4 polynucleotide kinase as desribed by Maxam and Gilbert (10).

Gel Electrophoresis Six to 8% acrylamide gels were used in which 1/40 bis-acrylamide and 12% glycerol were also added. The gels were in 50mM Tris-borate buffer (pH 8.3) at 200 volts for 20 hrs. After electrophoresis, the gels were soaked for 5 min in 5 ug/ml ethidium bromide solution and then destained in distilled water for 1 hr. The gels were then illuminated with a short-wave transilluminator and a MP-3 Polaroid camera was used to photograph the gels. For DNA sequencing, denaturing gels in which 20% degassed acrylamide and 7M pure urea were dissolved in 50mM Tris-borate along with 1/29 bis-acrylamide.

DNA Sequence Determination DNA fragments labeled at both 5' termini with ^{32}P were cleaved by a second restriction enzyme to give single-end-labeled fragments and were subjected to the chemical degradation reactions specific for A, G, C, C+T as described by Maxam and Gilbert (10). The partial cleavage products were run on a 20% denaturing gel in 50mM Tris-borate buffer at about 1,000 volts for three successive intervals. The ladder gel was then exposed to Dupont Cronex films at -70C with Kodak Hi-plus tungstate intensifying screens.

RESULTS AND DISCUSSIONS

Restriction Map of pL21-3 Plasmid

Digestion of pL21-3 with HpaII yielded two fragments of 1,240 and 910 base pairs(bp) not found in a parallel digest of the parent vector plasmid

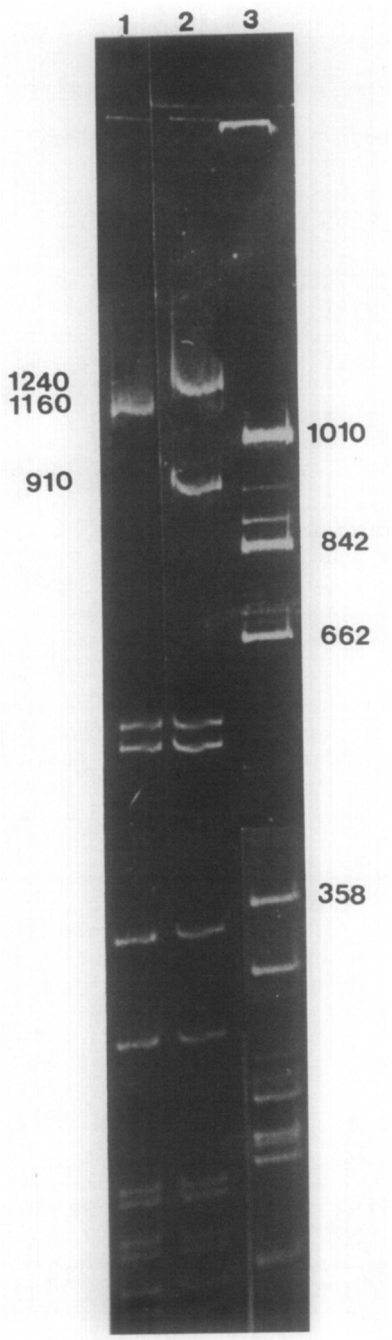


Fig.1. Estimation of size of kappa light chain sequence inserted in plasmid pMB9. pL21-3 (lane 2) recombinant plasmid DNA and pMB9 (lane 1) were digested with HpaII and electrophoresed on a 6% acrylamide gel in 50mM Tris-borate buffer (pH 8.3). Sizes were estimated by running along with ϕX174 RF DNA digested with AluI (lane 3). Digestion of pL21-3 by HpaII yielded two fragments of 1,240 and 910 base pairs not found in a parallel digest of the parent vector plasmid pMB9.

pMB9 (Fig.1). The size of the insert in pL21-3 was estimated to be (1,240 + 910) - 1,160 = 990 bp, including the poly-A tails. The was somewhat different from the 910 bp previously reported by Wall et al.(5) who used a 3% horizontal agarose gel compared to a 8% acrylamide gel used in our experiments.

Cleavage of the Hpall 1,240 bp fragment with Hinfl yielded three subfragments of 660, 430 and 145 bp (Table 1). However, there was no cleavage site for Hinfl on the 910 bp fragment. With the published data on the various restriction sites (relative to EcoRl site) it can be concluded that the pL21-3 sequence was inserted in a reverse orientation (see Fig. 3) in contrast to pL21-1 (8). The ordering of the two Hinfl sites (in the 1,240 bp fragment) was such that they could be placed at the immediate opposite sides of the 3' poly-A tails (Fig.3). Since the recent sequence data on MOPC21 mRNA by Hamyln et al. (9) have indicated a Hinfl site at the extreme end of the 3' untranslated mRNA, the insert sequence in pL21-3 should represent the entire amino acid sequence of the constant region and its 3' untranslated terminus, and possibly most, if not all the entire variable and its untranslated terminus. With the known sequence of the constant and the 3' untranslated regions of the MOPC21 mRNA (9), restriction of the 1,240 Hpall fragment with HaeIII, MbOII and Alu I (Table 1) resulted in unambiguous mapping of all those restriction sites (Fig. 3).

Table 1 Secondary cleavage of the 1,240 bp and 910 bp fragments generated by primary digestion of pL21-3 plasmid with Hpall. To help determine the order of the various subfragments, the 1,240 bp and 910 bp fragments were labeled with ^{32}P at both 5' termini and subjected to the secondary cleavage. The underlined subfragments represent those arising from the 5' ends after cleavage.

L3 Hpall fragment	Hinf I	HaeIII	MboII	AluI
1,240bp	660 430 145	800 460	760 210 200 90	780 260 80 60 50
910bp	NO SITES	310 165 150 92 80	—	310 302 165 88 76

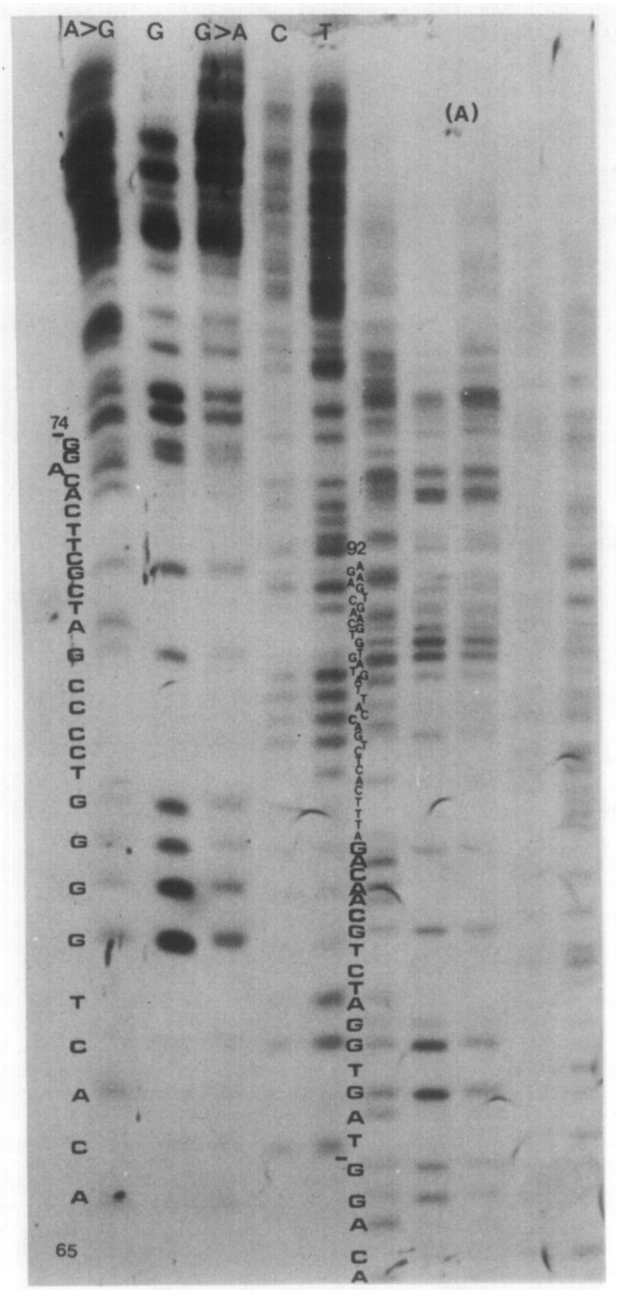
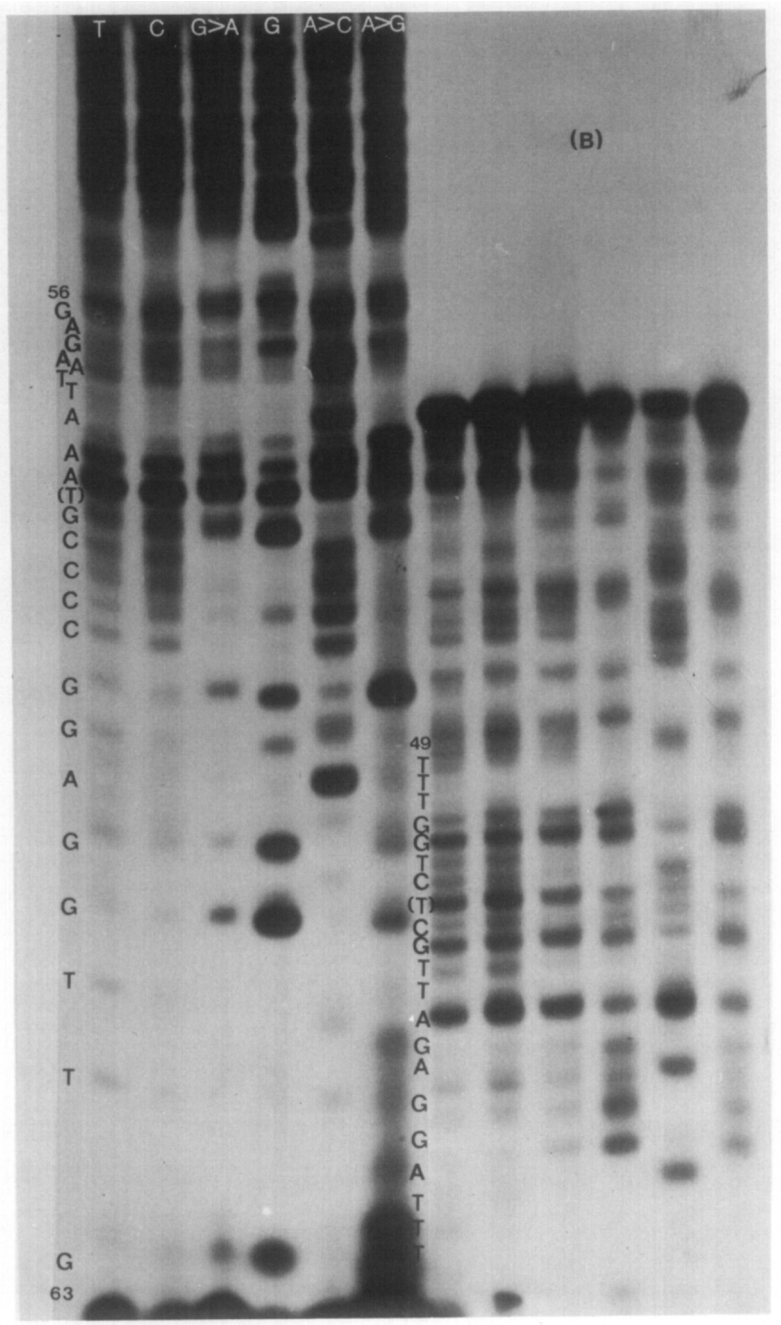


Fig.2. Autoradiogragh for determination of the nucleotide sequence from the internal HpaII site in pL21-3 toward the AluI and HaeIII sites respectively. DNA sequencing reactions (10) were performed on the 5'-labeled HpaII 1,240 bp and 910 bp fragments after cleavage with HinfI and HaeIII, respectively. A) The coding nucleotide sequence was read from the HpaII site downstream toward the AluI site and corresponds to amino acids 64 to 91 in the variable region of MOPC21



kappa light chain DNA (11). B) The nucleotide sequence was read from the HpaII site upstream toward the HaeIII site, and was in the anti-sense strand compared to that shown in Fig.3. The amino acids found to match were from 59-63 in the variable region. The two ambiguities that were bracketed to show these two nucleotides were not immediately apparent from the sequencing gel.

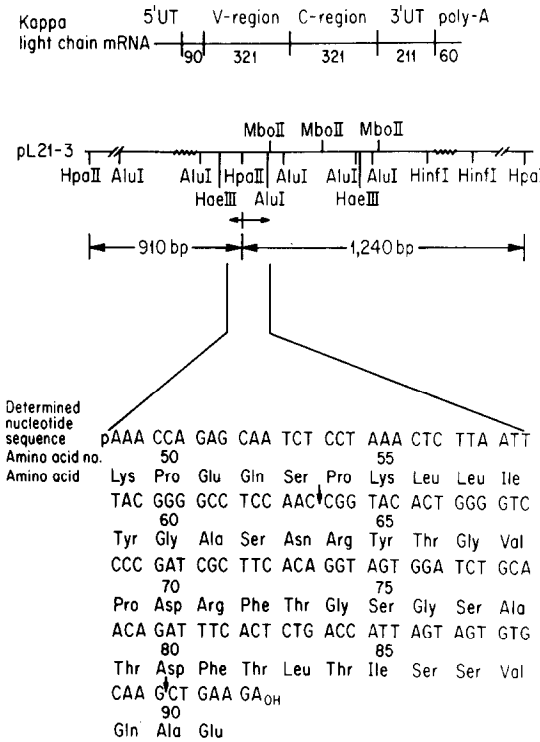


Fig.3. Restriction map of the MOPC21 immunoglobin kappa light chain gene sequences cloned in pMB9. The kappa light chain mRNA is divided into five regions, a 5' untranslated region, a putative leader sequence, a variable region sequence, a constant region sequence and an 3' untranslated region of 211 base pairs. The insert in pL21-3 is estimated to be 990 bp long, extending from the 5'leader region to the 3' untranslated region. Insert fragments were identified and sized along with digestion of pMB9 plasmid and ϕX174 RF DNA on a 6% acrylamide gel. Ordering of the insert fragments was done unambiguously by carrying out a series of secondary and double digestions. The determined nucleotide sequence (see Fig. 2) was from regions flanking the internal HpaII site. The sequence was found to correspond to the amino acids 49 to 91 in the variable region of the MOPC21 kappa light chain polypeptide.

When the 910 bp HpaII fragment was digested with AluI, five subfragments of the sizes, 310, 305, 148, 88 and 76 were produced (table 1). To help orient these five subfragments, the 910 bp fragment was labeled with gamma ^{32}P-ATP at both its 5' ends and then cleaved with AluI - only the 310 and 165 bp subfragments were seen in the autoradiograph. Since our sequence data have established the HpaII site to be exactly 159 bp from the 5' terminus of the variable region (see following section), it would appear that one of the AluI sites should be placed a few bp into the leader sequence region (Fig. 3), a

result consistent with the data reported for pL21-3 by Strathearn et al.(8). Cleavage of the 910 bp HpaII fragment with HaeIII resulted in a small 80 bp subfragment that agreed well with the predicted HaeIII site (85 bp from the HpaII site) from the amino acid sequence of MOPC21 kappa light chain determined by Svasti and Milstein (11). As a summary, the insert sequence in pL21-3 would start from the 5' leader sequence and extend all the way to the 3' noncoding terminus.'

DNA Sequence Determination

In order to unambiguously identify the relation of cloned sequence, a nucleotide sequence was determined after cleavage of the insert by restriction endonucleases. We have isolated the 1,240 bp HpaII fragment that encompasses the variable and constant regions (see Fig. 2), kinased with ^{32}P to label the 5' ends of this fragment and cleaved with HinfI to produce a 660 bp 5'-labeled subfragment. As shown in Fig.2A, 87 nucleotides, as determined by the chemical partial degradation reactions of Maxam and Gilbert (10), were found to correspond to the amino acids 64 to 91 previously reported by Svasti and Milstein (11). The first codon, GGC which was eluted off the ladder gel, had been verified by a separate experiment. The sequence confirmed the mapping of two adjacent restriction sites, AluI and MbOII which are about 75 bp downstream from the HpaII site. In addition, there had been ambiguity of assigning HpaII site to either amino acid 64 or amino acid 66 in the variable region by restriction mapping, as noted in pL21-1 by Strathearn et al. (8). Now, it can be concluded that the correct placement of HpaII site is at amino acid 64.

The nucleotide sequence flanking the 5' side of the HpaII site was also determined after cleavage of the 5' labeled 910 bp HpaII fragment with HaeIII. The determined nucleotide sequence (Fig. 2B) was from the anti-sense strand and correlated with the amino acids 49 to 63. There remain some ambiguities in the assignment of two nucleotides, which have been bracketed to show that these two nucleotides are not immediately apparent from the sequencing gel. Experiments in progress will clarify these ambiguities while sequence analysis is also presently underway to further characterize the entire variable region of the cloned cDNA.

A complete characterization of the cloned kappa light chain cDNA would provide pure hybridization probes for understanding immunoglobin gene structure and its expression in the mammalian genome. Uses of related kappa light chain and lambda light chain cDNA probes have recently produced exciting findings on the structure and organization of immunoglobin genes in mouse (12,13). In case of the pL21-3 clone, it is important that the putative leader region and the adjacent regions be sequenced in order to determine whether there is indeed a continguity between the leader and its adjacent regions. There are two lines of evidence that another cloned cDNA, pL21-4, may represent the 5' leader and non-coding sequences - no detectable hybridization between a pure kappa light chain constant region specific mRNA from MPC-11 cells and the insert in pL21-4 (5), and a HpaII site that is normally present in the variable regions of both pL21-1 and pL21-3 is absent from pL21-4 (unpublished data). Sequence analysis of the pL21-4 cloned sequence could shed light on the organization and expression of mRNA. Partial sequence analysis of the pL21-4 clone in this laboratory has indicated a high CT content (data not published), as is seen in the 3' noncoding region by Hamyln et al. (9). Therefore, characterization and isolation of the leader sequences and the 5' untranslated sequences as specific hybridization probes will help immensely in the understanding of the molecular organization of immunoglobin gene structure in mouse and other mammalian genome.

Acknowlegement This research was supported by American Cancer Society grant VC-240 and PHS grants CM18586, HL21831 and 78-CA2316

REFERENCES

1. Higuchi,R., Paddock,G., Wall,R., and Salser,W. (1976) Proc. Natl. Acad. Sci. 73, 3146-3150
2. Salser,W., Browne,J., Clarke,P., Heindell, H., Higuichi,R., Paddock,G., Roberts,J., Studnicka,G., and Zakar,P. (1976) Nucleic Acid Res. 19, 177-204
3. Maniatis, T., Kee,S.G., Efstratiadis,A., and Kafatos,F.C. (1976) Cell 8:163-182
4. Rabbits, T.H. (1976) Nature 260, 221-225
5. Wall,R., Gilmore-Hebet, Higuichi,R., Komaromy,M., Paddock,G., Strommer,J., and Salser,W. (1978) Nucleic Acid Res. 5, 3113-3128
6. Grunstein,M, and Hogness,D.S. (1975) Proc. Nat. Acad. Sci. 72, 3961-3965
7. Gilmore-Hebet,M., Hercules,K., Komaromy,M., and Wall,R. (1978) Proc. Nat. Acad. Sci. 75, 6044-6048

8. Strathearn,M.D., Strathearn,G., Akopiantz,P., Lui,A., Paddock,G., and Salser,W. (1978) Nucleic Acid Res. 5, 3101-3112
9. Hamyln,P.H., Brownlee,G.G., Cheng,C.C., Gait,M.J., and Milstein,C. (1978) Cell 15,1067-1075
10. Maxam,A, and Gilbert,W. (1977) Proc. Nat. Acad. Sci. 74,560-564
11. Svasti, J., and Milstein,C. (1972) Biochem. J. 128,427-444
12. Seidman,J.G., Leder,A., Edgell,M.H., Polsky,F., Tilghman,S.M., Tiemeier,D.C., and Leder, P. (1978) Proc. Nat. Acad. Sci. 75, 3881-3885
13. Tonegawa, S., Brack,C., Hozumi,N., and Schuller,R. (1977) Proc. Nat. Acad. Sci. 74, 3518-3522