

## Germ line transcription of the immunoglobulin heavy chain locus directs production of mu chain without VDJ.

J Schwaber, B Malone

*J Clin Invest.* 1992;89(6):2046-2052. <https://doi.org/10.1172/JCI115816>.

**Research Article**

Immunoglobulin VDJ recombination is associated with transcriptional activation of the Ig variable region elements. We have previously described a novel Ig mu chain protein and mRNA produced by pre-B cell hybrids from normal and X-linked agammaglobulinemic bone marrow. We have now characterized the mRNA encoding this protein and find that it is composed of a 5' leader sequence spliced to C mu (LS-C mu), lacking the variable (V), diversity (D), and joining (J) gene sequences. The leader sequence is encoded by a novel exon 16 kb upstream of the JH locus. Transcription of the germ line heavy chain locus from this LS exon results in transcriptional activation of the JH locus, apparently the initial step in commitment to B lymphoid development. Polymerase chain reaction amplification of normal bone marrow shows that these germ line LS-C mu transcripts are a product of bone marrow pre-B cells. Production of LS-C mu commences a sequential process of transcriptional activation, with concordant translation of Ig rearrangement intermediates, in the process of creating a productive VDJ rearrangement.

**Find the latest version:**

<https://jci.me/115816/pdf>



## Germ Line Transcription of the Immunoglobulin Heavy Chain Locus Directs Production of $\mu$ Chain without VDJ

Jerrold Schwaber and Barbara Malone

Department of Pathology, Harvard Medical School and Center for Blood Research, Boston, Massachusetts 02115; and Department of Pathology, Hahnemann University, Philadelphia, Pennsylvania 19102

### Abstract

Immunoglobulin VDJ recombination is associated with transcriptional activation of the Ig variable region elements. We have previously described a novel Ig  $\mu$  chain protein and mRNA produced by pre-B cell hybrids from normal and X-linked agammaglobulinemic bone marrow. We have now characterized the mRNA encoding this protein and find that it is composed of a 5' leader sequence spliced to  $C\mu$  (LS- $C\mu$ ), lacking the variable (V), diversity (D), and joining (J) gene sequences. The leader sequence is encoded by a novel exon 16 kb upstream of the  $J_H$  locus. Transcription of the germ line heavy chain locus from this LS exon results in transcriptional activation of the  $J_H$  locus, apparently the initial step in commitment to B lymphoid development. Polymerase chain reaction amplification of normal bone marrow shows that these germ line LS- $C\mu$  transcripts are a product of bone marrow pre-B cells. Production of LS- $C\mu$  commences a sequential process of transcriptional activation, with concordant translation of Ig rearrangement intermediates, in the process of creating a productive VDJ rearrangement. (*J. Clin. Invest.* 1992. 89:2046–2052.) Key words: immunoglobulin • recombination • B lymphocytes • leader sequence • pre-B cells

### Introduction

The genes encoding the variable regions of immunoglobulin are assembled in developing lymphocytes by selection and recombination of diverse gene elements. The Ig heavy chain V region is formed by recombination of variable ( $V_H$ ),<sup>1</sup> diversity (D), and joining (J) gene elements (1, 2) in an ordered rearrangement process (3). First a D gene element is recombined to a  $J_H$  segment, followed by recombination of a  $V_H$  to the formed  $DJ_H$  structure. The sequential steps of recombination are catalyzed by a single recombinase (4). The steps of recombination are

associated with transcriptional activation of the recombining gene elements, suggesting that the recombinase is directed in its sequential steps by control of access (5, 6). In  $V_H$  to  $DJ_H$  recombination, the  $DJ_H$  intermediate is transcribed from promoters upstream of several of the D gene elements (7, 8), and unrearranged  $V_H$  elements are transcribed from the promoters required for mature  $V_HDJ_HC\mu$  transcription. Transcription of the unrearranged  $J_H$  locus has been reported (6), but the transcripts and promoter that directs this transcription have not been characterized.

The human antibody deficiency disease X-linked agammaglobulinemia (XLA) results from failure of B lymphocyte development (9). Patients with the major phenotype of XLA have normal numbers of pre-B cells in bone marrow, but only limited numbers of mature B lymphocytes in peripheral circulation (10). We have previously proposed that the arrest in B lymphoid development in XLA results from failure of Ig V(D)J recombination (11). This proposal was based on the identification of truncated  $\mu$  chains produced as mRNA and translated polypeptides in pre-B cell hybrids derived from the bone marrow of patients with XLA. This novel transcript was identified in some pre-B cell hybrids from normal fetal liver, but was the only product identified in hybrids from three patients with XLA.

We have now characterized these truncated  $\mu$  chains, and report that they are novel transcripts of the unrearranged Ig heavy (H) chain locus, representing transcriptional activation of the  $J_H$  locus. The transcripts are composed of a 5' leader sequence spliced to the  $\mu$  constant region (LS- $C\mu$ ), lacking the V, D, or J sequence characteristic of the previously described  $D\mu$  ( $DJ_H$  rearrangement) or mature  $\mu$  ( $VDJ_C\mu$ ). LS- $C\mu$  is transcribed from an exon 16 kb upstream of the  $J_H$  locus in the unrearranged H chain locus. Transcription from this exon entails transcriptional activation of the unrearranged  $J_H$  elements. By polymerase chain reaction (PCR) amplification, we have identified LS- $C\mu$  transcripts in normal bone marrow, further indicating that it represents a normal product of VDJ recombination.

### Methods

The cell hybrids are as described previously (11, 12). Normal human bone marrow was obtained by iliac crest biopsy from a normal volunteer, with aspiration into preservative-free heparin as anticoagulant. Written informed consent was obtained following procedures reviewed and approved by the Committee on Clinical Investigation, Children's Hospital, Boston. Mononuclear cells were separated by buoyant density centrifugation on Ficoll-paque (Pharmacia Fine Chemicals, Piscataway, NJ), and adherent cells were removed by adherence to plastic petri dishes overnight. For nucleic acid preparation, cells were pelleted by centrifugation, the medium aspirated, and stored at  $-80^\circ$  until used to make RNA or DNA.

Address correspondence to Jerrold Schwaber, Ph.D., Department of Pathology, MS 435, Hahnemann University, Broad and Vine, Philadelphia, PA 19102.

Received for publication 8 November 1991 and in revised form 17 January 1992.

1. Abbreviations used in this paper: D, diversity; H, heavy; J, joining; LS, leader sequence; LS- $C\mu$ , 5' LS spliced to the  $\mu$  constant region; PCR, polymerase chain reaction;  $V_H$ , variable; XLA, X-linked agammaglobulinemia.

*J. Clin. Invest.*

© The American Society for Clinical Investigation, Inc.

0021-9738/92/06/2046/07 \$2.00

Volume 89, June 1992, 2046–2052

RNA was isolated by guanidine-HCl lysis as described by Chirgwin et al. (13), followed by differential ethanol precipitation to remove DNA. PolyA-containing sequences were enriched by chromatography in oligo-dT cellulose. cDNA was synthesized with cloned Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories, Gaithersburg, MD) using oligo-dT or random hexamers (Promega Corp., Madison, WI) for priming. A cDNA library was constructed from oligo-dT-primed cDNA by annealing dC-tailed cDNA into dG-tailed pBR322. The insert from the  $C_{\mu}$  plasmid  $p\mu$ HM2 and a 61-bp EcoRI-Hinfi fragment at the 5' end of  $C_{\mu}$  (nt 47-108, see [9]) were used to screen the library. The inserts of  $p\mu$ HM2 and  $\mu$ b9 were later subcloned into pUC18. LS- $C_{\mu}$  sequence was determined by both Maxam and Gilbert's (14) and dideoxy methods (15), and the sequence was determined for both strands. Genomic DNA was extracted, digested with restriction endonucleases, and Southern blotted to nitrocellulose (Schleicher & Schuell, Inc., Keene, NH) as described by Maniatis et al. (16). 5' terminal cDNA probes from  $\mu$ b9 and 73 2-33 were EcoRI fragments of cDNA clones isolated by electrophoresis through a 5% acrylamide gel (14). EcoRI sites are present at the 5' terminus of these cDNAs (added by adapters for cloning) and at nucleotide 47 of the  $C_{\mu}$  sequence.

Polymerase chain reaction amplifications were performed in a thermal cycler (Perkin-Elmer Cetus Instruments, Norwalk, CT) with buffers, dNTPs, and TaqI polymerase from Perkin-Elmer Cetus. cDNA for PCR was synthesized with random primers, and incorporation of [ $^{32}$ P]dCTP in an aliquot used to quantitate the cDNA. Primers were synthesized by Dr. A. Whitehead (Harvard Medical School, Boston MA), using a synthesizer (Applied Biosystems, Inc., Foster City, CA). The primers used were two nested primers for  $C_{\mu}$ :

CMI 5' TGATGGAGTCGGGAAGGAAGTCCTGT 93-127 of  $C_{\mu}$   
 CMIII 5' TGCTGGACTTTGCACACCACGTGTTCTGT 247-278 of  $C_{\mu}$ .

These are reverse complement sequences of the coding sequence of  $C_{\mu}$ .

LSP 5' TTCCATTCGGTGATCAGCACTGAACACAGAGGACTC -84 to -48 of  $\mu$ b9.

Three primers from  $V_H$  gene 5' flanking leader sequence 1st exons were:

61RILS 5' TCACCAGAGCTCCAGACAAT -158 to -138 of 6-1R1 ( $V_H6$ )

19IILS 5' TACTTTCTGAGAGTCCTGGACC -173 to -153 of 1-9II ( $V_H4$ )

212LS 5' CCTTCTCTACAGAAGCCT -181 to -163 of 21-2 ( $V_H1$ ).

These primers were derived from sequences published by Berman et al. (17). A map of the Ig H chain locus, and the positions of these primers is given in Fig. 1. The positions listed are nucleotides upstream from the 5' end of the FR1 sequence, which places them in comparable positions to LSP with respect to the leader sequence first exon of these  $V_H$  gene elements. cDNA was first amplified with CMIII as 3' primer. If no product was detected in an agarose gel, 1  $\mu$ l of a 1:20 dilution of the primary PCR was reamplified with either CMIII or CMI nested primer.

The conditions for PCR amplification were 94° (denaturation) for 1 min, 70° (annealing) for 1 min, 72° (reaction) for 3 min, for 30 cycles. The annealing temperature was reduced to 65° for the LSP/CMIII reaction with normal bone marrow. Precautions were taken to prevent cross-contamination.

Primers to amplify a fragment entirely within  $C_{\mu}$  were used as positive control for amplification. These primers:

CM4R  
 5' CAACTTGGTGGACTTGTTTCAGGAAGATGCT +700 to +730 reverse complement of  $C_{\mu}$  coding strand and CM5F  
 5' AAGACGTGCTCTTCCAGTGATTG +299 to +323 of  $C_{\mu}$  coding

yielded a 433-bp amplification product. We used these primers as control in the following way: All reagents for amplification of an experimental cDNA except primers were mixed in a single tube. These reagents included the test cDNA and TaqI polymerase. 98- $\mu$ l aliquots were placed in individual tubes and the relevant primers (e.g., 61RILS and CMIII, or CM4R and CM5F) were added, followed immediately by thermal cycling for amplification. The CM4R/CM5F reaction then served as a test for addition of all necessary reagents for the amplification reaction and provided a control for comparison of amplification between experiments run at different times.

Amplification products were identified by Southern blotting of samples in agarose gels. The DNA was transferred either to nitrocellulose as previously described (16) or to Sure Blot Hybridization Membrane (Oncor Inc., Gaithersburg, MD) following the manufacturer's instructions. Blots were probed with the insert  $C_{\mu}$  fragment from the  $p\mu$ HM2 plasmid.

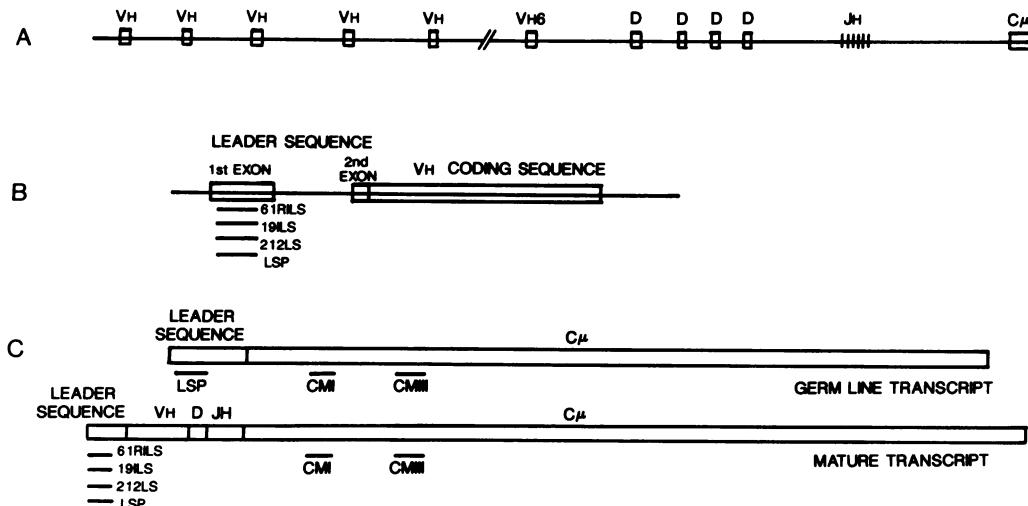


Figure 1. Map of primers used for PCR examination of Ig transcripts. (A) The germ line Ig heavy chain locus, not drawn to scale, is shown.  $V_H$  gene elements are carried in the genome over a large region starting ~ 90 kb upstream of the  $C_{\mu}$  coding sequence. The single element  $V_H6$  gene, named 6-1RI by Berman et al. (17) is the most 3' of the  $V_H$  gene elements. The locations of the  $V_H4$  (1-9II) and  $V_H1$  (21-2) gene elements, and the  $V_H$  gene elements that have near identity to the  $\mu$ b9 (LSP primer) leader sequence are not known. The LS exon is

not shown. (B) The structure of  $V_H$  gene elements is shown. A leader sequence first exon is separated from the remainder of the leader sequence (second exon) and  $V_H$  coding sequence by an intron of ~ 100 bp. The primers 61RILS, 19IILS, and 212LS were taken from sequences of the first leader sequence exons of the 6-1RI, 1-9II, and 21-2  $V_H$  gene elements. These primers are analogous in position to the LSP primer on the 56P1  $V_H$  gene element. (C) The positions of the leader sequence and  $C_{\mu}$  primers on LS- $C_{\mu}$  germ line and  $V_HDJ_H C_{\mu}$  mature transcripts are shown. The LSP primer yielded PCR amplification products with both germ line and mature transcripts (see Fig. 2 and Results). The 61RILS, 19IILS, and 212LS primers yielded PCR amplification product only with mature transcripts (see Results).

Phage clones spanning the J<sub>H</sub> to C<sub>μ</sub> intron were CH 4-22 (a twin to CH 4-38 [18]) and CH 28-6, provided by P. Leder (Harvard Medical School), were as described by Ravetch et al. (18) and clone C7p6.2 was as described by Rabbitts et al. (19). Cosmids C17p1, c17p3, and Cos<sub>μ</sub>6, provided by H. Schroeder (University of Alabama at Birmingham) and D. Cox (Hospital for Sick Children, Toronto), were as described by Schroeder et al. (20).

## Results

**cDNA sequence encoding the truncated  $\mu$  chain.** Four pre-B cell hybrids were previously described (11). LSM-T, LSM-A, and LSM-SB21 resulted from fusion of mononuclear cells from bone marrow of three different patients with X-linked agammaglobulinemia. LSH6 was a clone derived from fusion with normal fetal liver pre-B cells. This clone was selected for study because it produced a truncated  $\mu$  chain that comigrated with the  $\mu$  chain produced by the XLA hybrids. The LSM 2.7 human myeloma used as fusion partner made no detectable Ig as protein or RNA, and had deleted the C<sub>μ</sub> gene (12, and unpublished results). The four hybrids (LSM-T, LSM-A, LSM-SB21, and LSH6) were previously characterized to produce truncated  $\mu$  heavy chain without variable region as both mRNA and polypeptide.  $\mu$  chain proteins were reduced in size to 54 and 57 kD, compared with 66 and 69 kD for normal  $\mu$  chain (secretory and membrane forms). mRNA-encoding  $\mu$  chain was reduced in size, due to absence of 300 nt of 5' terminal sequence encoding the V region.

To determine the 5' terminus of the truncated  $\mu$  chains, a cDNA library was constructed by oligo-dT priming of mRNA from the LSH6 hybrid. Two clones encoding  $\mu$  chain were isolated. Restriction mapping of the clones confirmed that there were no deletions of the C<sub>μ</sub> sequence. The sequence of the clones was determined from the EcoRI site at nucleotide 47 of C<sub>μ</sub>. The sequence lacks V<sub>H</sub>, D, or J<sub>H</sub> gene sequence of mature  $\mu$  chain or previously characterized immature D $\mu$ . In place of this sequence, one clone ( $\mu$ b9) has an 88-nt sequence that has near identity to the first LS exon of two human V<sub>H</sub> gene elements (Fig. 2) (21). The second clone has an identical sequence but extends only 63 nt upstream of C<sub>μ</sub>. The putative LS has an

initiation codon (ATG) 61 nt upstream of the 5' terminus of C<sub>μ</sub> and an open reading frame downstream. The ATG is in phase for translation of C<sub>μ</sub>. Compared with the product of primer-extension experiments, the longer clone is missing 65 nt of the full message, presumably 5' untranslated sequence.

Because of the size of the cDNA sequence, its open reading frame, and the absence of other detectable  $\mu$  chain message, it seems likely that this truncated message encodes the 54/57 kD protein identified in the pre-B cell hybrids. Since leader sequences are removed from the nascent polypeptide with passage into the lumen of the endoplasmic reticulum, this truncated polypeptide is C<sub>μ</sub>, and the transcript is LS-C<sub>μ</sub>. This novel  $\mu$  chain is the product of the pre-B parents of these hybrid cells. The myeloma parent of these hybrids could not have contributed the novel  $\mu$  chains because it lacks the C<sub>μ</sub> gene by Southern blotting and produces no detectable  $\mu$  chain RNA, either by Northern blotting or by PCR amplification (unpublished results).

**LS-C<sub>μ</sub> is a product of normal pre-B cells.** Immune fluorescence experiments previously suggested that some pre-B cells from normal bone marrow produced  $\mu$  chain without V region antigen (11). To determine whether LS-C<sub>μ</sub> is a product of (these) normal pre-B cells, we examined normal bone marrow directly by PCR amplification. cDNA prepared from total RNA of normal human bone marrow was amplified with primers derived from C<sub>μ</sub> and LS. A DNA fragment of 360 bp, the size predicted for the 5' terminus of LS-C<sub>μ</sub> with these primers, was identified (Fig. 3). This product was isolated from a gel, reamplified, and cloned into pUC18. The sequence of this molecule has near identity to the LS cDNA clone from the LSH6 pre-B cell hybrid (Fig. 3 B). The three nucleotides that differ are conservative for known leader sequence functions. The differences might result from an artifact of PCR amplification. However, isolation of seven independently derived PCR clones with this sequence from this sample of normal bone marrow leads us to suggest that this sequence represents polymorphism in the LS. Truncated  $\mu$  chains composed of LS-C<sub>μ</sub> are likely candidates for the product of normal bone marrow pre-B cells previously identified to produce  $\mu$  chain without V region antigen.

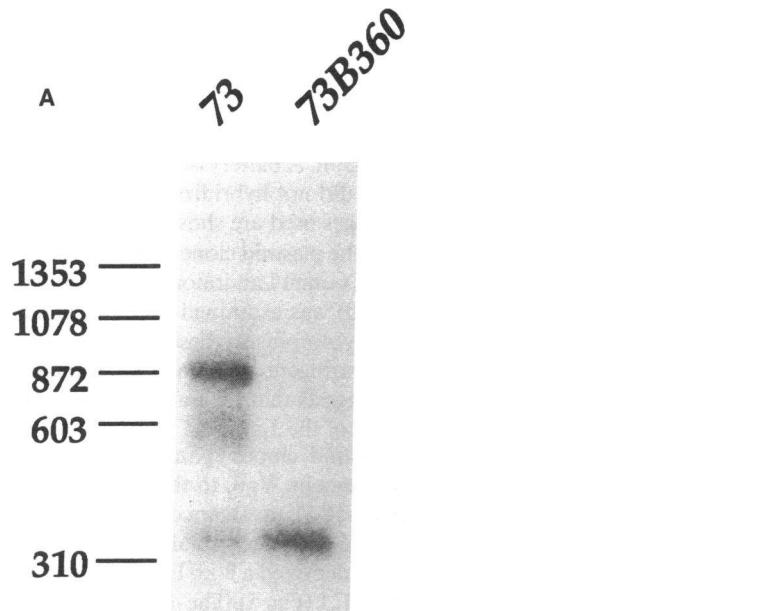
A

		-80		-70		-60		-50											
		/		/		/		/											
$\mu$ b9	AGT	TTC	CAT	TCT	GTG	ATC	AGC	ACT	GAA	CAC	AGA	GGA	CTC	ACC	ATG				
56P1	---	---	---	---	---	-T-	---	---	---	---	---	---	---	---	---				
		-40		-30		-20		-10							-1+1				
		/		/		/		/							/				
$\mu$ b9	GAG	TTT	GGG	CTG	ACC	TGG	GTT	TTC	CTC	GTT	GCT	CTT	TTA	A	GA	C <sub>μ</sub>			
56P1	---	---	---	---	-G-	---	---	---	---	---	---	---	---	---	---	intron			

B

$\mu$ B9	Met	Glu	Phe	Gly	<u>Leu</u>	Thr	Trp	Val	Phe	Leu	Val	Ala	Ile	Leu	Lys
56P1	-	-	-	-	-	Ser	-	-	-	-	-	-	-	-	-

**Figure 2.** Sequence of cDNA clone  $\mu$ b9 from the pre-B hybrid LSH6. An ATG initiation codon is located at position -43, identical in location to the initiation codon for the known leader sequence first exon of the V<sub>H</sub> gene element 56P1 cloned as cDNA from fetal liver (21). 56P1 is spliced to the second exon of the V<sub>H</sub> gene element, containing the remainder of the leader sequence and the coding sequence of the variable region.  $\mu$ b9 is spliced to the 5' terminus of C<sub>μ</sub>, providing the initial A nucleotide to the 5' terminal GA, necessary for in frame translation of C<sub>μ</sub>. There are two single base differences between  $\mu$ b9 and 56P1. The single nucleotide change in the coding region of this proposed leader sequence results in conservative substitution of Thr for Ser. The  $\mu$ b9 sequence has been submitted to GenBank under accession number M81847.



B

		-50		-40		-30		-20				
		/		/		/		/				
$\mu$ b9	primer	GGA	CTC	ACC	ATG	GAG	TTT	GGG	CTG	ACC	TGG	GTT
73 2-33	primer	---	---	---	---	---	---	---	---	---	---	---

			-10							
		/								
$\mu$ B9		TTC	CTC	GTT	GCT	CTT	TTA	A	GA	$C\mu$
73 2-33		---	---	---	---	---	---	---	---	$C\mu$

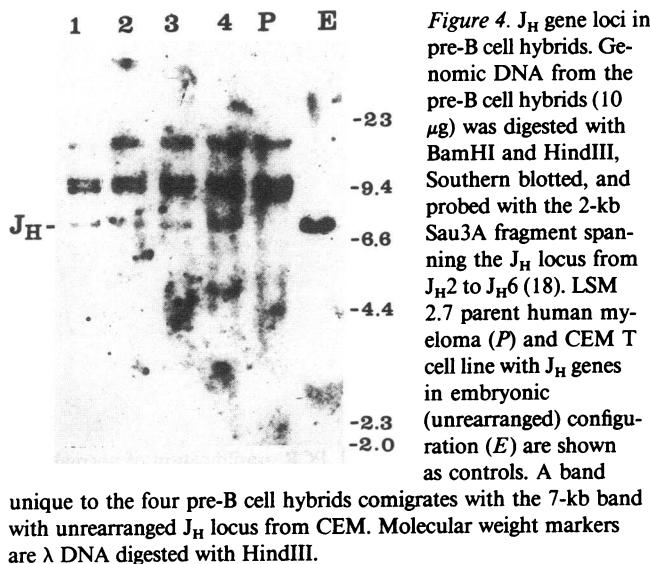
C

$\mu$ B9	Met	Glu	Phe	Gly	Leu	Thr	Trp	Val	Phe	Leu	Val	Ala	Leu	Leu	Arg
73 2-33	-	-	Leu	-	-	Cys	-	-	-	-	-	-	-	-	-

**Figure 3.** PCR amplification of normal bone marrow cDNA. (A) Autoradiogram of Southern blot of the PCR products. The first lane contains the product of amplification of cDNA from normal bone marrow with the LSP/CMIII primers (sample 73). The 360-bp band was excised from a low melt agarose gel, and reamplified with the same primers (lane 2, 73B360). The 360-bp product from the secondary amplification was isolated from low melt agarose and cloned into pUC18 after addition of EcoRI adapters. (B) Sequence of the 360-bp cDNA 73 2-33 cloned from normal bone marrow in comparison to  $\mu$ b9. Three nucleotides differ from the sequence of  $\mu$ b9, causing two amino acid substitutions in the proposed leader sequence. Seven independently derived clones from this experiment had the 73 2-33 sequence. At the 3' end of the sequence A is contributed by the LS while the GA are contributed by  $C\mu$ , forming an asymmetric splice characteristic of immunoglobulin gene family members. The 73 2-33 sequence has been submitted to GenBank under accession number M81848. (C) Derived amino acid sequence from clone 73 2-33. Substitution of Leu for Phe is conservative. Substitution of Cys for Thr, while not conservative, occurs in the leader sequences of other  $V_H$  gene elements (see Kabat et al. [38]).

*LS-C $\mu$  is transcribed before recombination.* Southern blotting of the  $J_H$  locus from the pre-B cell hybrids suggests that LS-C $\mu$  results from transcriptional activation of the Ig heavy chain locus before recombination. HindIII/BamHI double digestion was used to separate the  $J_H$  locus from  $C\mu$ . Genomic DNA from the four pre-B cell hybrids contained a 7-kb band that comigrates with the embryonic (unrearranged)  $J_H$  locus band from an unrearranged T cell line (Fig. 4). This band is present against a background of multiple rearranged  $J_H$  bands that are contributed by the myeloma parent of the hybrids, and is the only band unique to the hybrid cells that are shared in all four hybrids. All of the other  $J_H$  bands in the pre-B cell hybrids are also present in the myeloma parental genomic DNA. We suggest that LS-C $\mu$  is transcribed from the unrearranged H chain locus.

*Does LS-C $\mu$  result from transcriptional activation of  $V_H$  gene elements?*  $C\mu$  with a leader sequence nearly identical to the leader sequence first exon of some  $V_H$  gene elements might result from transcriptional activation of the entire H chain locus, including the  $V_H$  gene elements, or from transcription from a specific LS exon upstream of the  $C\mu$  locus. For LS-C $\mu$  to result from transcriptional activation of  $V_H$  genes, long transcripts spanning more than 100 kb from the unrearranged  $V_H$  gene elements to  $C\mu$  would utilize the splice sequence at the 3' end of some or all  $V_H$  gene element first exons, yielding truncated  $\mu$  chains with leader sequence first exons of diverse origin. We tested for truncated  $\mu$  chains with alternate  $V_H$  leader sequences by PCR amplification of LSH6 pre-B cell hybrid cDNA with primers specific for leader sequence first exon of several  $V_H$  gene elements. Primers used were specific for  $V_H6$



**Figure 4.**  $J_H$  gene loci in pre-B cell hybrids. Genomic DNA from the pre-B cell hybrids (10  $\mu$ g) was digested with BamHI and HindIII, Southern blotted, and probed with the 2-kb Sau3A fragment spanning the  $J_H$  locus from  $J_{H2}$  to  $J_{H6}$  (18). LSM 2.7 parent human myeloma (*P*) and CEM T cell line with  $J_H$  genes in embryonic (unrearranged) configuration (*E*) are shown as controls. A band

unique to the four pre-B cell hybrids comigrates with the 7-kb band with unrearranged  $J_H$  locus from CEM. Molecular weight markers are  $\lambda$  DNA digested with HindIII.

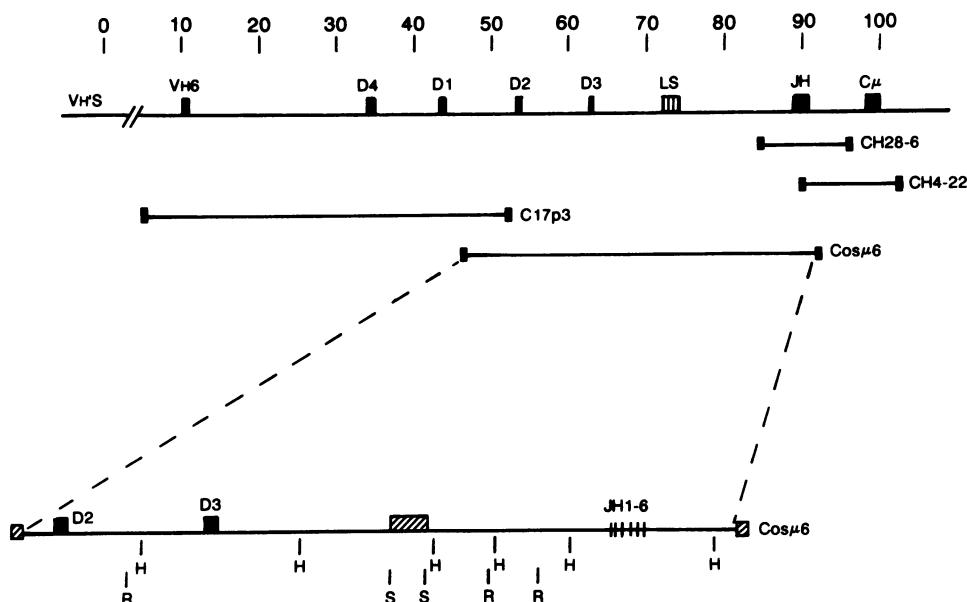
(single element family that is most proximal to  $C_\mu$ ) and 1-9II ( $V_{H4}$  gene family) and 21-2 ( $V_{H1}$  gene family). Amplification product was identified only with the LSP primer derived from the  $\mu b9$  sequence (data not shown), indicating that LS- $C_\mu$  results from transcription from a specific LS exon.

**Mapping of the LS exon.** The absence of transcriptional activation of  $V_H$  gene elements suggests that LS- $C_\mu$  results from transcription of a specific LS exon upstream of  $C_\mu$ . Direct cloning of the LS exon has been hindered by the near identity of  $\mu b9$  with the leader sequence first exons of some  $V_H$  gene elements, and the apparent polymorphism of the LS between different individuals. As an alternate approach, we have assumed that the LS exon lies between the  $V_{H6}$  gene element, the most 3'  $V_H$  gene element, and  $C_\mu$ , and restriction mapped a collection of plasmid, phage, and cosmids spanning this region.

Nelson et al. (22) identified transcriptional activation of the  $C_\mu$  locus arising from the  $J_H$  to  $C_\mu$  intron. Termed sterile  $\mu$  transcripts, they have not been found to be processed into mRNA or to yield a sequence that can be translated. To rule out LS- $C_\mu$  transcription arising from this intron, we examined phage clones that span this region. A battery of LS probes, both oligonucleotides and cDNA, did not hybridize with Southern blots of these clones (the phages used are shown in Fig. 5, the negative blot is not shown). The plasmid clone C7p6.2 from T. Rabbitts (Medical Research Council Laboratory for Molecular Biology, Cambridge, England) was examined also, but is not shown (19). To allow for polymorphism, these hybridizations were carried out at reduced stringency (42° and 6 $\times$  standard saline citrate [SSC]). This suggests that LS transcription must arise from an exon upstream of the  $J_H$  locus.

Southern blotting of cosmid clones spanning the region from the most 3'  $V_H$  gene element,  $V_{H6}$ , to the  $J_H$  locus indicates that the LS exon lies 16–18 kb upstream of  $J_{H1}$ . Examination of the cosmid clone *cos $\mu$ 6* with 5' terminal fragments from the  $\mu b9$  and 73 2-33 cDNAs identified a 9-kb HindIII fragment that is 14–23 kb upstream of  $J_{H1}$  (Fig. 5). The  $\mu b9$  cDNA probe was removed at temperatures above 48° and salt below 2 $\times$  SSC, while the 73 2-33 probe was resistant to melting to 55° and 0.1 $\times$  SSC, further evidence for polymorphism in the LS. Digestion of *cos $\mu$ 6* with additional enzymes refined this map, locating the LS exon to a 2.3-kb *SacI* fragment that lies 14–16 kb upstream of  $J_{H1}$ . No additional fragments that hybridize with LS probes were found on the other cosmid clones extending to the  $V_{H6}$  gene element (Fig. 5). This location for the LS exon is exactly analogous to the  $\kappa^o$  promoter upstream of the  $J_\kappa$  locus (23), and suggests that the LS exon serves as a transcriptional promoter for the  $J_H$  locus.

Transcription from the LS exon yields a 26–28-kb primary transcript with splice donor sites at the 3' end of the LS and at the 3' ends of each  $J_H$  element, with a single splice acceptor site at the 5' end of  $C_\mu$ . LS- $C_\mu$  mRNA was only identified from



**Figure 5.** Mapping of the LS exon.

The phage and cosmid clones used to map the LS exon are shown below the map of the region from  $V_{H6}$  to  $C_\mu$  on chromosome 14. The phage clones CH 28-6 and CH 4-22 (twin to CH4-38), as described by Ravetch et al. (18), showed no hybridization with LS probes, either oligonucleotides, or 5' *EcoRI* restriction fragments from the cDNA clones  $\mu b9$  and 73 2-33, when stringency was reduced to 42° at 2 $\times$  SSC. The cosmid clones, as described by Schroeder et al. (20) showed hybridization of both oligonucleotides and cDNA restriction fragments with the 9.0-kb HindIII band shown on the map. The oligonucleotides melted from this band above 42° at 2 $\times$  SSC. The cDNA restriction fragment from  $\mu b9$  was stable to 48° at 2 $\times$  SSC while the cDNA fragment

from 73 2-33 was stable to 0.1 $\times$  SSC at 55°. *SacI* digestion refines this map, identifying a 2.3-kb fragment at the 3' end of the 9.0-kb HindIII fragment. Upstream in this map is to the left, with the  $V_{H6}$  gene element 77 kb 5' to the  $J_H$  gene elements. The scale above the top line is in kb. The map locations of the gene elements shown are derived from Schroeder et al. (20).

splicing of the LS splice donor site to the  $C\mu$  splice acceptor site. This product is translated into polypeptide. Alternate splicing from the  $J_H$  element splice donor sites would yield 16–17-kb RNAs that likely could not be translated. These alternate products would not have been identified by PCR amplification.

## Discussion

We have described a novel Ig H chain transcript that arises from an exon upstream of the  $J_H$  locus before recombination. LS- $C\mu$  production may represent the initial step in recombinational access for the unrearranged H chain locus, directing transcriptional activation of the  $J_H$  elements before  $DJ_H$  recombination. The recombinase has been proposed to be directed in its sequential steps of catalysis by control of access (24). This proposal is based on identification of transcription of the  $DJ_H$  intermediate and unrearranged  $V_H$  gene elements in Abelson virus-transformed murine pre-B cells (5). Transcriptional activation of the light chain  $J\kappa$  locus has been described, and  $V\kappa$  to  $J\kappa$  recombination is associated with transcriptional activation from the  $\kappa^\circ$  exon (6, 25). The LS exon serves as an analogous promoter for the H chain, directing transcriptional activation of the unrearranged  $J_H$  locus. LS- $C\mu$  transcripts are processed into mRNA and translated into polypeptide.

The first three steps in Ig recombination are associated with transcriptional activation. The unrearranged  $J_H$  locus is transcribed from the LS exon before D to  $J_H$  recombination (6). Unrearranged  $V_H$  gene elements and the  $DJ_H$  intermediate step of recombination are transcribed during  $V_H$  to  $DJ_H$  recombination (5, 6). The murine  $\kappa$  light chain locus has a leader sequence exon,  $\kappa^\circ$ , analogous to the human H chain LS exon (23, 25).  $\kappa^\circ$  has near identity to the first leader sequence exon of  $V\kappa$  gene elements, and lies 3–4 kb upstream of the  $J\kappa$  locus. Transcripts from the  $\kappa^\circ$  exon are processed into truncated  $\kappa$  chain mRNA. In addition to the Ig locus, T cell antigen receptor gene elements likely are transcribed.  $D\beta$  elements direct transcription and likely translation of the  $D\beta J\beta$  intermediate (26). Transcripts of  $T\alpha$  and  $T\beta$  have been described in human B cell precursor cell lines (27), and appear to be transcribed in T cells as well (28). Transcriptional activation of gene elements during recombination of Ig and T cell antigen receptors appears to be a rule.

mRNA transcripts of the same size as LS- $C\mu$  have been identified previously in murine cell lines, but not further characterized (29, 30). Kemp et al. (29) found a 1.9-kb  $C\mu$  transcript as the only Ig present in T cell and myeloid cell lines with unrearranged Ig loci, including the T cell line EL4. Schlissel et al. (6) described a  $\mu 0$  PCR amplification product that resulted from transcription of the unrearranged  $J_H$  elements in EL4. The PCR based assay used by Schlissel et al. (6) would not identify mRNA derived from an LS exon that was processed as LS- $C\mu$ . However, alternative processing of LS- $C\mu$  primary transcripts, using the splice donor sites at the 3' ends of each  $J_H$  locus, would yield RNAs that amplify as  $\mu 0$ . We suggest that there is a murine analogue of LS- $C\mu$  and the LS exon, leaving open the question of how alternative processing of these transcripts is regulated.

A mechanical interpretation of the access model entails sequential activation, first of the  $J_H$  locus, then of D gene elements for D to  $J_H$  recombination, then of  $V_H$  gene elements for  $V_H$  to  $DJ_H$  recombination. The linear arrangement of the gene elements does not fit with a strict interpretation of this model.

One human D gene element is located within the  $J_H$  gene locus (18), and several D gene elements are upstream of the most 3'  $V_H$  gene element (31). An alternate explanation of the observed transcriptional activation is simply that the entire locus opens up in preparation for recombination, transcription occurs coincidentally, and the sequential steps of recombination are regulated by a secondary mechanism. This view is consistent with a report of differential regulation for D to  $J_H$  versus  $V_H$  to  $DJ_H$  recombination in murine transfection vectors (32). In this view, conservation of transcriptional promoters upstream of the  $J_H$  and  $J\kappa$  loci, and of the D gene elements is not strictly required for recombination. Instead, they could then be viewed either as evolutionary relics, or could be of import for translation of the intermediate structures into protein.

The significance of translation of these immature products into protein is not clear. LS- $C\mu$  transcripts are clearly translated into truncated  $\mu$  chains by the pre-B cell hybrids previously described and likely represent the 3–5% of pre-B cells from normal bone marrow previously identified to be  $C\mu + V_H^-$  by immune fluorescence (9). Cell-free translation of LS- $C\mu$  from the pre-B cell hybrids showed that some of the mRNA includes the 3'  $\mu$ mem exon necessary for insertion in the cell membrane, raising the possibility of these molecules posing as membrane Ig.  $D\mu$  intermediates are translated into protein in both murine and human cells (7, 8, 33). Again, some of these molecules include the  $\mu$ mem exon, and are known to be expressed as membrane Ig. The truncated  $\mu$  chains, both LS- $C\mu$  and  $D\mu$ , may combine with the light chain surrogates lambda5 and Vpre-B to form a pseudo-immunoglobulin molecule without variable region (34, 35). Polypeptides with sequence predicted by the  $\kappa^\circ$  nucleic acid sequence have been reported (36, 37). The  $D\beta$  transcripts of T cell antigen receptor have an upstream ATG initiation codon with following open reading frame with the properties of a leader sequence, suggesting that this intermediate may be translated as well (26). The conservation of transcription and translation of the several steps of recombination in both murine and human cells is the strongest argument for a regulatory role of these intermediate products.

We have previously proposed that the failure of B cell development in patients with XLA results from a failure of H chain VDJ recombination (11). VDJ recombination is the primary event in pre-B cells. LS- $C\mu$  was the only Ig product found in uncloned pre-B cell hybrids from three XLA patients. In contrast, only 3 of 15 pre-B cell hybrids derived from normal fetal liver produced LS- $C\mu$  (unpublished results). Determination of whether LS- $C\mu$  is the predominant product of XLA bone marrow pre-B cells, by direct examination of the products of XLA bone marrow pre-B cells, will be reported in a later publication.

## Acknowledgments

We gratefully acknowledge the education in molecular biology provided by Stuart Orkin in furtherance of these studies. We thank B. Kim for providing the normal bone marrow sample, S. Goff and S. Orkin for sequencing  $\mu b9$ , P. Leder for the phage clones spanning  $J_H$  to  $C\mu$ , and the plasmid clone with  $J_H$  probe, T. Rabbitts for the plasmid clone spanning  $C\mu$ , D. Cox, M. Hofker, and H. Schroeder for the cosmid clones, and S. Orkin, E. Hoffman, and J. Sylvester for critical reading of the manuscript. F. Rosen provided valuable discussion. In memory of Sam Latt, who provided critical intellectual support for this research.

Supported by National Institutes of Health grants AI21163 and 21165. J. Schwaber was a Scholar of the American Heart Association.

## References

1. Sakano, H., Y. Kurosawa, M. Weigert, and S. Tonegawa. 1981. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy chain genes. *Nature (Lond.)* 290:562-565.
2. Early, P., H. Huang, M. Davis, K. Calame, and L. Hood. 1980. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA, V<sub>H</sub>, D, and J<sub>H</sub>. *Cell* 19:981-992.
3. Alt, F., G. Yancopoulos, T. Blackwell, C. Wood, E. Thomas, M. Boss, R. Coffman, N. Rosenberg, S. Tonegawa, and D. Baltimore. 1984. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1209-1219.
4. Yancopoulos, G., T. K. Blackwell, and F. Alt. 1986. Introduced T cell receptor variable region gene segments recombine in pre-B cells: evidence that B and T cells use a common recombinase. *Cell* 44:251-259.
5. Blackwell, T. K., M. Moore, G. Yancopoulos, H. Suh, S. Lutzker, E. Selsing, and F. Alt. 1986. Recombination between immunoglobulin variable region gene segments is enhanced by transcription. *Nature (Lond.)* 324:585-589.
6. Schlissel, M., L. Corcoran, and D. Baltimore. 1991. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. *J. Exp. Med.* 173:711-720.
7. Reth, M., and F. Alt. 1984. Novel immunoglobulin heavy chains are produced from DJ<sub>H</sub> gene segment rearrangements in lymphoid cells. *Nature (Lond.)* 312:418-423.
8. Schwaber, J., and R. Chen. 1988. Premature termination of variable region rearrangement in B lymphocytes from X-linked agammaglobulinemia. *J. Clin. Invest.* 81:2004-2009.
9. Rosen, F., M. Cooper, and R. Wedgwood. 1984. The primary immunodeficiencies. *N. Engl. J. Med.* 311:235-242.
10. Pearl, E., L. Vogler, A. Okos, W. Crist, A. Lawton, and M. Cooper. 1978. B lymphocyte precursors in human bone marrow: an analysis of normal individuals and patients with antibody-deficiency states. *J. Immunol.* 120:1169-1175.
11. Schwaber, J., H. Molgaard, S. Orkin, H. Gould, and F. Rosen. 1983. Pre-B cells from normal and X-linked agammaglobulinemic individuals produce  $\mu$  chains without associated variable regions. *Nature (Lond.)* 304:355-358.
12. Schwaber, J., M. Posner, S. Schlossman, and H. Lazarus. 1984. Human-human hybrids secreting pneumococcal antibodies. *Hum. Immunol.* 9:137-143.
13. Chirgwin, J., A. Przybyla, R. MacDonald, and R. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-5300.
14. Maxam, A., and W. Gilbert. 1980. Sequencing end labeled DNA with base specific cleavages. *Methods Enzymol.* 65:449-560.
15. Sanger, F., S. Nicklen, and A. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5466.
16. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
17. Berman, J., S. Mellis, R. Pollock, C. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C. Cantor, and F. Alt. 1988. Content and organization of the human Ig V<sub>H</sub> locus: definition of three new V<sub>H</sub> families and linkage to the Ig CH locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:727-738.
18. Ravetch, J., U. Siebenlist, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Structure of human immunoglobulin  $\mu$  locus: characterization of embryonic and rearranged J and D genes. *Cell* 27:583-591.
19. Rabbitts, T., A. Forster, and C. Milstein. 1981. Human immunoglobulin heavy chain genes: evolutionary comparisons of C $\mu$ , C $\delta$ , and C $\gamma$  and associated switch regions. *Nucleic Acids Res.* 9:4509-4524.
20. Schroeder, H., M. Walter, M. Hofker, A. Ebens, K. Willems van Dijk, L. Liao, D. Cox, E. Milner, and R. Perlmutter. 1988. Physical linkage of a human immunoglobulin heavy chain variable region gene segment to diversity and joining region elements. *Proc. Natl. Acad. Sci. USA.* 85:8196-8200.
21. Schroeder, H., J. Hillson, and R. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science (Wash. DC)* 238:791-793.
22. Nelson, K., J. Haimovich, and R. P. Perry. 1983. Characterization of productive and sterile transcripts from the immunoglobulin heavy chain locus: processing of  $\mu$ m and  $\mu$ s mRNA. *Mol. Cell. Biol.* 3:1317-1332.
23. Van Ness, B., M. Weigert, C. Coleclough, E. Mather, D. Kelley, and R. Perry. 1981. Transcription of the unrearranged mouse C $\kappa$  locus: sequence of the initiation region and comparison of activity with a rearranged V $\kappa$ -C $\kappa$  gene. *Cell* 27:593-602.
24. Alt, F., T. K. Blackwell, R. dePinho, M. Reth, and G. Yancopoulos. 1986. Regulation of genome rearrangement events during lymphocyte differentiation. *Immunol. Rev.* 89:5-30.
25. Nelson, K., D. Kelley, and R. P. Perry. 1985. Inducible transcription of the unrearranged  $\kappa$  constant region locus is a common feature of pre-B cells and does not require DNA or protein synthesis. *Proc. Natl. Acad. Sci. USA.* 82:5305-5309.
26. Siu, G., M. Kronenberg, E. Strauss, R. Haars, T. Mak, and L. Hood. 1984. The structure, rearrangement and expression of D $\beta$  gene segments of the murine T cell antigen receptor. *Nature (Lond.)* 311:344-350.
27. Calman, A. F., and B. M. Peterlin. 1986. Expression of T cell receptor genes in human B cells. *J. Exp. Med.* 164:1940-1957.
28. Loh, E., J. Elliott, S. Cwirla, L. Lanier, and M. M. Davis. 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science (Wash. DC)* 243:217-220.
29. Kemp, D. J., A. Harris, S. Cory, and J. Adams. 1980. Expression of the immunoglobulin C $\mu$  gene in mouse T and B lymphoid and myeloid lines. *Proc. Natl. Acad. Sci. USA.* 77:2876-2880.
30. Alt, F., N. Rosenberg, V. Enea, E. Siden, and D. Baltimore. 1982. Multiple immunoglobulin heavy chain gene transcripts in Abelson murine leukemia virus transformed lymphoid cell lines. *Mol. Cell. Biol.* 2:386-400.
31. Matsuda, F., E. Shin, Y. Hirabayashi, H. Magaoka, M. Yoshida, S. Zong, and T. Honjo. 1990. Organization of variable region segments of the human immunoglobulin heavy chain: duplication of the D5 cluster within the locus and interchromosomal translocation of variable region segments. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:2501-2506.
32. Ferrier, P., B. Krippel, T. K. Blackwell, A. Furley, H. Suh, A. Winoto, W. Cook, L. Hood, F. Constantini, and F. Alt. 1990. Separate elements control DJ and VDJ rearrangement in a transgenic recombination substrate. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:117-125.
33. Nickerson, K., J. Berman, E. Glickman, L. Chess, and F. Alt. 1989. Early human IgH gene assembly in Epstein-Barr virus transformed fetal B cell lines. *J. Exp. Med.* 169:1391-1403.
34. Tsubata, T., and M. Reth. 1990. The products of pre-B cell specific genes (lambda5 and VpreB) and the immunoglobulin mu chain form a complex that is transported onto the cell surface. *J. Exp. Med.* 172:973-976.
35. Kurasuyama, H., A. Kudo, and F. Melchers. 1990. The proteins encoded by the VpreB and lambda5 preB cell specific genes can associate with each other and with mu heavy chain. *J. Exp. Med.* 172:969-972.
36. Burstein, Y., R. Zemell, F. Kantor, and I. Schechter. 1977. Independent expression of the gene coding for the constant domain of immunoglobulin light chain: evidence from sequence analyses of the precursor of the constant region polypeptide. *Proc. Natl. Acad. Sci.* 74:3157-3161.
37. Rose, S., W. M. Kuehl, and G. Smith. 1977. Cloned MPC11 myeloma cells express two kappa genes: a gene for a complete light chain and a gene for a constant region polypeptide. *Cell* 12:453-462.
38. Kabat, E., T. Wu, M. Reid-Miller, H. Perry, and K. Gottesman. 1987. *Sequences of proteins of immunological interest*. U. S. Department of Health and Human Services Publication. pp 11-19.