Secondary genomic rearrangement events in pre-B cells: V_HDJ_H replacement by a LINE-1 sequence and directed class switching

George D. Yancopoulos, Ronald A. DePinho, Kathryn A. Zimmerman, Stuart G. Lutzker, Naomi Rosenberg¹ and Frederick W. Alt

Department of Biochemistry and Institute for Cancer Research, Columbia University College of Physicians and Surgeons, New York, NY 10032, and ¹Cancer Research Institute and Department of Pathology, Tufts University Medical School, 136 Harrison Avenue, Boston, MA 02111, USA

Communicated by B. Pernis

We describe rearrangement events which alter expression from a productive V_HDJ_H rearrangement in an Abelson murine leukemia virus-transformed pre-B cell line. One such rearrangement results in replacement of the initially expressed variable region gene by a site-specific join between the open reading frame of a LINE-1 repetitive element and a remaining J_H segment. We discuss this event in the context of the 'accessibility' model of recombinase control, and with respect to similar rearrangements involved in oncogene activation. In another subclone of the same pre-B cell line, altered heavy chain expression resulted from a μ to $\gamma 2b$ class switch recombination which occurred by a recombination–deletion mechanism but involved a complex inversion. We provide evidence that the germline $\gamma 2b$ region is specifically expressed in pre-B cell lines and early in normal development. We propose that the predisposition of pre-B cell lines to switch to $\gamma 2b$ production may reflect a normal physiological phenomenon in which the switch event is directed by an increased 'accessibility' of the germline $\gamma 2b$ locus to switch–recombination enzymatic machinery. Our findings support the hypothesis that the apparently distinct recombination systems involved in variable region gene assembly and heavy chain class switching are both directed by the accessibility of their substrate gene segments.

Key words: accessibility/class switching/germline transcription/LINE-1 repeats/recombination

Introduction

Expression of immunoglobulin (Ig) heavy (H) chain genes is dependent on two distinct types of somatic recombination events (for review, see Alt *et al.*, 1986). The first results in the formation of a complete H-chain variable region gene from three widely separated germline elements: variable (V_H), diversity (D) and joining (J_H) segments (Earley *et al.*, 1980; Sakano *et al.*, 1980). Site-specific rearrangement of these segments is apparently mediated by highly conserved recognition sequences which consist of a characteristic heptamer and an A/T-rich nonamer separated by a spacer region (Earley *et al.*, 1980; Sakano *et al.*, 1980). The joining events occur in an ordered sequence during pre-B cell differentiation: D to J_H joins occur first and on both chromosomes, followed by the appendage of an upstream V_H region to form a complete V_HDJ_H variable region gene (Alt *et al.*, 1984). Although the potential exists for a B cell to assemble two different variable region genes on its two alleles, any given B

cell clone produces a single H chain: this phenomenon is referred to as H-chain allelic exclusion (Pernis *et al.*, 1965). H-chain allelic exclusion has been suggested to occur by regulation of the joining process: current evidence suggests that the H-chain protein expressed from the first productive (H-chain producing) V_HDJ_H rearrangement made by a cell signals for the cessation of further V_H to DJ_H rearrangement in that cell and its progeny (Alt *et al.*, 1981, 1982a, 1984; Weaver *et al.*, 1985; Rusconi and Kohler, 1985). Recent experiments indicate that a single site-specific recombinase performs all of the many variable region gene assembly events seen in both B and T cells, and that the tissue and stage specificity of these joining events is mediated by varying the 'accessibility' of the component gene segments to the common recombinase (Yancopoulos and Alt, 1985; Yancopoulos *et al.*, 1986).

A second type of recombination event, referred to as H-chain class switching, allows the progeny of a given B cell to maintain the parental H-chain antigen binding specificity (encoded by the variable region gene) while varying its effector function [encoded by different constant region (C_H) genes] (reviewed in Shimizu and Honjo, 1984; Marcu and Cooper, 1982). The order of the C_H genes has been determined to be 5'(V_HDJ_H)– C_μ – C_δ – $C_{\gamma 3}$ – $C_{\gamma 1}$ – $C_{\gamma 2b}$ – $C_{\gamma 2a}$ – C_ϵ – C_α (Shimizu *et al.*, 1982). Long stretches of repeated sequences (S regions), which lie upstream to each C_H gene, apparently mediate the switch event (reviewed in Shimizu and Honjo, 1984). Recombination between S_μ and downstream S regions deletes intervening sequences and places the assembled variable region gene just upstream of the new C_H gene (subsequently referred to as the 'classical recombination/deletion mechanism').

Initial analyses of the genetic events involved in both variable region gene assembly and H-chain class switching were based on comparison of the H-chain gene structure in the germline with that in terminally differentiated B-cell tumors. Abelson murine leukemia virus (A-MuLV)-transformed pre-B cell lines have provided dynamic representations of early B cell development. Several of these lines perform V_H to DJ_H joining during propagation in culture (reviewed by Yancopoulos and Alt, 1986; Alt *et al.*, 1986). In addition, A-MuLV pre-B cell lines can undergo H-chain class switch events by the classical recombination/deletion mechanism (Alt *et al.*, 1982b; Burrows *et al.*, 1983; Akira *et al.*, 1983; DePinho *et al.*, 1984). Notably, all such reported events involved a switch from μ to $\gamma 2b$ production; switching to other isotypes has not been described in A-MuLV-transformed pre-B cell lines.

We previously suggested that secondary rearrangement events involving V_H to V_HDJ_H recombination (designated $V[H]$ replacement) might provide a mechanism for diversifying the initially restricted H-chain variable region repertoire (Yancopoulos *et al.*, 1984). To further explore this possibility we examined subclones of A-MuLV-transformed pre-B cell lines which exhibited altered H-chain protein expression from initial V_HDJ_H rearrangements which, in the parent line, had yielded normal μ H-chain production. Here we demonstrate that in the 300-18 line (see Alt

Fig. 1. Heavy chain gene organization in PD subclones. Approximately 10 μ g of genomic DNA from the indicated lines was digested with *Eco*RI, fractionated by electrophoresis through 1.0% agarose gels, and blotted to nitrocellulose. Duplicate filters were probed with either the J_H probe (panel A) or the 3' $S_{\gamma 2b}$ probe (panel B). Sizes are indicated in kilobases in parentheses.

et al., 1981, 1984 for characterization of 300-18; this line has also been referred to as 'PD' in Lewis *et al.*, 1982, and will subsequently be referred to as 'PD' for simplicity) such altered H-chain protein expression results from secondary rearrangements in the vicinity of the initially expressed V_HDJ_H complex. Although we did not find the anticipated secondary V_H -replacement events, our analyses of two novel secondary rearrangements in the PD line may provide insight into the mechanism and control of the two major types of recombination events (variable region gene assembly and H-chain class switching) which occur during B cell differentiation.

Results

Secondary J_H -associated rearrangement events

Several subclones of the PD line were found to have lost normal μ H-chain production. Two were chosen for extensive molecular analyses: subclone PD31 had replaced μ H-chain production with lower levels of a larger Ig-cross-reactive protein (Lewis *et al.*, 1982), while PD6 had replaced μ production with $\gamma 2b$ production (not shown). The four germline J_H segments are contained within a single 6.2-kb *Eco*RI fragment (Figure 1, lane 1). V_HDJ_H (or DJ_H) formation replaces the 5' end of this germline fragment and results in an *Eco*RI fragment that is usually different in size from the germline fragment; Southern blotting analysis using a J_H probe allows identification of the rearranged J_H -hybridizing *Eco*RI fragments in a given line (Alt *et al.*, 1984). The parental PD line, which is a μ -producing pre-B cell line (Lewis *et al.*, 1982), had two rearranged J_H -hybridizing *Eco*RI fragments previously identified as V_HDJ_H rearrangements (Alt *et al.*, 1984).

Fig. 2. Partial restriction maps of primary and secondary rearrangements at the productive 'A' allele. (A) Secondary 'A2' rearrangement in PD31 subclone; VRS refers to the V-replacement sequence. (B) Productive V_HDJ_H rearrangement in PD parent line, referred to as 'A'. (C) Secondary 'A1' rearrangement in PD6 which results from productive class switch event. The region of A1 which matches the germline $S_{\gamma 2b}$ sequence presented in panel D is denoted by a bold line; the dashed line in A1 indicates the region which does not match either the A rearrangement presented in panel B or the germline $S_{\gamma 2b}$ region presented in panel D below; arrows denote the inverted $S_{\gamma 2b}$ sequences in this region which might derive from the R_3 repeat presented in panel D as indicated. (D) Germline $S_{\gamma 2b}$ region in both BALB/c mice (Kataoka *et al.*, 1981) and in the PD parent line, which is derived from NIH/Swiss strain. The two groups of repeats reported by Wu *et al.* (1984) are denoted by the arrows marked R_1 and R_2 ; these groups of repeats may extend further upstream than has been reported, and this is indicated by dashed region of the arrows representing these repeats. R_3 denotes a possible third repeat in the uncharacterized upstream region of $S_{\gamma 2b}$. (E) and (F) Schematic representation of mechanism accounting for the structure of the switch region in 'A1' rearrangement. Arrows indicate the germline orientation of the $S_{\gamma 2b}$ region, and 'D' denotes the small deletions in the $S_{\gamma 2b}$ region which were found in the A1 rearrangement. The circled 'E' in panels A–D refers to the heavy chain transcriptional enhancer element. Restriction sites: R, *Eco*RI; B, *Bam*HI; H, *Hind*III; Bg, *Bgl*II; Bst, *Bst*EII; P, *Pst*I; X, *Xba*I; N, *Nco*I; S, *Sac*I.

The 5.5-kb fragment and the allele it represents is designated 'A' and the 3.0-kb fragment and the allele it represents is designated 'B' (see Figure 1, lane 2). Comparison of the J_H -associated rearrangement patterns in subclones (designated PD6, PD15 and PD31) of the PD line allowed us to define secondary rearrangements and also to determine which of the parental rearrangements encoded the productive rearrangement. The PD15 subclone had lost the 3.0-kb 'B' fragment and maintained the parental 5.5-kb 'A' fragment unchanged (Figure 1, lane 3); because this line maintains normal μ protein production (not shown), we could define the 'A' allele as the productively rearranged allele in the parental line. The remaining two subclones (PD6 and PD31) retained the non-productive 'B' allele in the parental form, but each had further rearranged the productive 'A' allele (the altered 'A' fragment in PD6 was designated 'A1' and the altered 'A' fragment in PD31 was designated 'A2', Figure 1, lanes 4 and 5). Thus, the changes in protein production by the PD subclones

Fig. 3. Nucleotide sequence of relevant regions of the rearrangements presented in Figure 2. (A) Sequences of VRS-J_{H3} join and comparison with germline J_{H3} sequence and various LINE-1 sequences. The VRS-J_{H3} sequence is presented in triplets which are in-phase with the J_{H3} coding region. The VRS-J_{H3} sequence was determined downstream from the *EcoRI* site and upstream from the *PstI* site which border the VRS, as indicated in Figure 2A. The VRS-J_{H3} sequence is compared in its entirety with the homologous region of a complete mouse LINE-1 sequence (nucleotides 3197–3595 of the L1Md-A2 sequence in Loeb *et al.*, 1986) and in its latter portion with a partial mouse LINE-1 sequence (nucleotides 42–122 of L3 fragment, R3 portion in Gebhard *et al.*, 1972), a human LINE-1 sequence (nucleotides 869–939 of *HindIII* repeat in Manuelidis, 1982) and the germline J_{H3} sequence (Sakano *et al.*, 1980). The various recognition heptamers and nonamers are indicated; the putative LINE recognition sequences are compared with the murine V_{L2} recognition sequences (Bernard *et al.*, 1978). (B) Sequence near the switch site in A1 rearrangement. The switch site (compared with the corresponding region of the germline J_{H3}-C_{H3} intron which is numbered according to Gillies *et al.*, 1983) and the following stretch of inverted S_{2b} sequence (marked by solid underline) is indicated; these sequences are followed by normally oriented S_{2b} sequences which can be aligned with the germline S_{2b} region in NIH/Swiss and BALB/c mice (numbered according to Wu *et al.*, 1984); a deletion in the A1 rearrangement with respect to the germline sequences is indicated. At the inversion point the germline sequence no longer matches the sequence of the A1 rearrangement. (C) The inverted S_{2b} sequence, marked by solid underline in panel B is inverted so that it can now be aligned to both a downstream portion of the A1 rearrangement itself (which was presented in panel B) and to the BALB/c germline region. A deletion in the inverted S_{2b} sequence with respect to the sequences with which it is aligned is indicated.

were correlated with secondary rearrangements of the productive A allele. To understand the alterations of the A allele in detail, both parental J_H-hybridizing *Eco*RI fragments and the two altered A1 and A2 *Eco*RI fragments were isolated by molecular cloning into the lambda phage vector Ch16A.

Loss of normal μ protein production in PD31 is due to a recognition sequence-mediated recombinational event involving a LINE-1 repetitive element

The restriction maps of the A and A2 *Eco*RI fragments diverged near the unrearranged J_H3 segment in the A fragment, suggesting that loss of normal μ protein production in the PD31 subclone was due to deletion of the productive V_H-DJ_H2 complex present on the parental A allele by a rearrangement to the downstream J_H3 segment (Figure 2B and A). Furthermore, a V probe specific to the variable region expressed from the parental A allele (denoted V_A) did not hybridize to either the A2 rearrangement or to any RNA sequences in PD31 (data not shown). To further understand the nature of the apparent deletion event, the nucleotide sequence of the A2 fragment near the divergence point was determined. This sequence analysis revealed that a new element (designated VRS for V replacement sequence) had joined site-specifically to J_H3 (Figure 3A). A probe prepared from the VRS displayed a highly repetitive nature when used to analyze DNA and RNA blots and thus could not be used to isolate the germline counterpart of the rearranged VRS. However, comparison of the VRS sequence with known repetitive elements revealed 94% homology to a region within a previously sequenced full-length mouse LINE-1 repetitive element (designated L1Md-A2 by Loeb *et al.*, 1986; see Figure 3A). The LINE-1 repeats comprise the major family of long interspersed repetitive elements in the mouse genome; homologous families are present in most mammals, including humans (Singer and Skowronski, 1985). Both primate and mouse LINEs contain long open reading frames (designated 'ORFs'), apparently conserved during evolution, which are similar in size and organization to the genomes of retroviruses and eucaryotic transposable elements (Loeb *et al.*, 1986). The major ORF in LINEs displays homology to reverse transcriptase (Loeb *et al.*, 1986; Hattori *et al.*, 1986), supporting the notion that LINEs are transposable elements which encode proteins involved in their own dispersal; however, no LINEs-encoded protein product has ever been described. Like the related L1Md-A2 sequence, the VRS sequence may also derive from a full-length and functional LINE element; it also maintains an ORF in the region sequenced (Figure 3A). In fact, the VRS-J_H3 rearrangement has juxtaposed the 5' portion of the major ORF of a LINE-1 element with the J_H3-C _{μ} coding region, leaving both coding regions in the same continuous ORF (Figure 3A). Thus, the large Ig-cross-reactive protein which has replaced normal μ protein production in PD31 (see above) may be a fusion protein encoded by both LINE and Ig sequences.

Notably, the sequence of L1Md-A2 corresponding to the region just downstream of the recombination point in the VRS reveals potential recombination recognition sequences very similar to the recognition sequences bordering Ig gene segments: six of the seven base pairs in the potential heptamer and eight of the nine base pairs in the potential nonamer are identical to those found in the V λ 2 recognition sequences (Bernard *et al.*, 1978), and these potential recognition sequences contain seven of the 10 base pairs recently proposed as a 3' 'consensus core recognition sequence' by Rathbun and Tucker (1986) (Figure 3A). Although the putative L1Md-A1 recognition sequences are separated by a 15-bp spacer region, another partial LINE-1 sequence (Gebhard *et al.*, 1972)

has similarly situated potential recognition sequences separated by a 12-bp spacer (Figure 3A). Remarkably, a consensus Ig heptamer is located at this position in a human LINE sequence (Figure 3A). The precise location of the VRS-J_H3 join with respect to the J_H3 and presumed VRS recombination recognition sequences (Figure 3A), together with the complementarity of the corresponding 23- and 12-bp spacer regions, strongly suggest that this rearrangement is mediated by the same site-specific recombination mechanism as V_H to DJ_H joining.

Switching occurs by a recombination-deletion event resulting in an inverted S γ 2b region

To determine whether secondary rearrangement of the A fragment in the PD6 subclone was related to the switch to γ 2b production, a Southern blot identical to the one used to identify the J_H-associated rearrangements was screened with a probe specific for the S γ 2b region. The germline S γ 2b region is contained within a 6.8-kb *Eco*RI fragment; a switch event into the S γ 2b region will usually result in a non-germline sized *Eco*RI fragment that retains hybridization to an S γ 2b probe. Except for the PD6 subclone, all the lines examined had only the germline S γ 2b *Eco*RI fragment (Figure 1B). The PD6 line displayed a novel *Eco*RI fragment which hybridized with the S γ 2b probe, and this fragment co-migrated with the A1 *Eco*RI fragment defined above (compare lanes 4 and 9 in Figure 1A and B). The novel A1 *Eco*RI fragment in PD6 resulted from a rearrangement event which brought the J_H and S γ 2b regions into close proximity on the productively rearranged allele. Direct analysis of the molecularly cloned A1 *Eco*RI fragment verified the linkage of these two regions (see below), while comparison of the cloned A1 fragment with its genomic partner in PD6 demonstrated that no further rearrangement had occurred during molecular cloning (not shown). Analysis of poly(A)⁺ RNA, prepared from the PD subclones and assayed for hybridization to the V_A probe, revealed that the same variable region expressed on μ -sized RNA in the

Fig. 4. Truncated and normal γ 2b transcripts in pre-B cell lines and in normal murine tissues. 2 μ g of poly(A)⁺ RNA from the indicated cell lines, 4.0 μ g of poly(A)⁺ RNA from newborn (NB) murine liver, or 3.0 μ g of total RNA from adult murine spleens was fractionated on agarose-formaldehyde gels, transferred to nitrocellulose filters and assayed for hybridization to the indicated probes. Representative data are shown: multiple experiments on several different neonatal and adult tissue samples provided similar results; of the 15 pre-B cell lines screened, no detectable γ 2b transcripts were found in the lines 298-8, 22D7, 230-37, 220-8, 1-8 and 38B9, low but detectable levels were found in 22D6 and 38B7, while readily detectable levels were found in PD, 2M3, 18-48, 3-1, 40E4, 40E3 and 18-81A20. The characteristics of these lines (i.e. derivation, status of rearrangement at the H-chain locus and expression of μ protein) have been described in Alt *et al.* (1984).

PD parent line (not shown) and PD15 (Figure 4D) was expressed on a $\gamma 2b$ -sized RNA in PD6 (Figure 4D), verifying that switching had occurred at the A allele.

Nucleotide sequence analysis of the molecularly cloned A1 fragment allowed a detailed elucidation of its structure. The switch event occurred just downstream from the μ H-chain enhancer, which remained intact and was followed immediately by a 350-bp stretch of inverted $S_{\gamma 2b}$ sequence and then by normal $S_{\gamma 2b}$ sequence (Figure 3B,C). Comparison of the restriction map and nucleotide sequence of the A1 rearrangement with that of the previously published germline BALB/c $S_{\gamma 2b}$ region (Kataoka *et al.*, 1981) revealed that both were almost identical downstream from the inversion point in A1 (Figures 2C,D, 3B). However, inverted $S_{\gamma 2b}$ sequences are not found upstream from this position in the germline BALB/c $S_{\gamma 2b}$ region (Kataoka *et al.*, 1981). We molecularly cloned and characterized the corresponding region of the germline $S_{\gamma 2b}$ region from the PD parent line (derived from the NIH/Swiss strain), verifying that it was almost identical to the BALB/c $S_{\gamma 2b}$ region and did not contain an inversion (Figures 2D, 3B). Thus it appears that the inverted $S_{\gamma 2b}$ sequences in the A1 rearrangement arose during the switch event. A schematic representation of this join, and the least complex mechanism which we could envision to account for its structure, is presented in Figure 2E and F. The recombination/deletion/inversion events depicted involve two long almost identical sets of direct repeats (denoted R1 and R1 in Figure 2) within the $S_{\gamma 2b}$ region (Wu *et al.*, 1984); the joining model presented assumes that yet a third nearly identical set of repeats occurs (designated R3), at least in the germline of the NIH/Swiss mouse. A long stretch of uncharacterized $S_{\gamma 2b}$ sequences which could easily harbor this third set of repeats extends several kilobases upstream from the 5' *Xba*I site depicted in Figure 2D (Kataoka *et al.*, 1981). Several sequence mismatches between the inverted A1 $S_{\gamma 2b}$ region and R1 (Figure 3C) or R2 (not shown) support the notion that this region arose from a germline-encoded third set of repeats. Two small deletion events which occur near the switch site are also indicated (Figures 2F, 3B,C); deletions within the $S_{\gamma 2b}$ region have previously been noted, but were not analyzed at the sequence level (Kataoka *et al.*, 1981). The deletions characterized here appear to be mediated by homologous recombination; they maintain the 49-bp repeat structure of the $S_{\gamma 2b}$ region (Figure 3B,C).

Truncated $S_{\gamma 2b}$ transcripts in pre-B cells: directed switching?

The component gene segments involved in variable region gene assembly appear to become transcriptionally active just at or prior to the cell stage in which they are normally rearranged (reviewed in Yancopoulos and Alt, 1986). This transcription apparently reflects an increased 'accessibility' of these segments to recombinational machinery; evidence exists that regulating the 'accessibility' of these segments allows the targeting of a common recombinase to form tissue- and stage-specific joins (Yancopoulos *et al.*, 1986). We reasoned that a common switch recombinase might be directed to form class-specific switches by regulation of the accessibility of the various switch regions. The frequent switching of A-MuLV-transformed cell lines to $\gamma 2b$ production (for review, see Alt *et al.*, 1986; DePinho *et al.*, 1986) might be due to a specific increase in the 'accessibility' of the $\gamma 2b$ region in these cells. To further explore this possibility, we examined 15 independent A-MuLV-transformed pre-B cell lines (including PD) to determine whether specific transcriptional activation of the germline $\gamma 2b$ locus precedes its rearrangement; the lines chosen for analysis had no detectable DNA rearrangements in-

volving the $S_{\gamma 2b}$ region. Nine of the 15 pre-B cell lines examined expressed detectable levels of $C_{\gamma 2b}$ -hybridizing transcripts of which six had substantial levels (see legend to Figure 4A; representative data shown in Figure 4A). The $\gamma 2b$ transcripts in this set of A-MuLV transformants were slightly smaller than those encoding normal $\gamma 2b$ transcripts (compare lanes 1–5 with lane 6, and lane 7 with lane 8, in Figure 4). We analyzed the short $C_{\gamma 2b}$ transcripts in subclones of the PD line to determine if they contained variable region sequences. The μ -producing PD15 line contains only the initial A allele, has lost the non-productive B allele and maintains an unrearranged $\gamma 2b$ region (see above); yet PD15 contains levels of the short $C_{\gamma 2b}$ transcripts which are as high as the levels of normal-sized $\gamma 2b$ RNA in PD6, which has switched to $\gamma 2b$ (Figure 4C). Furthermore, the V_A probe identifies only μ RNA in PD15 and normal $\gamma 2b$ RNA in PD6 (Figure 4B) while a probe specific to the B allele does not identify $\gamma 2b$ -sized RNA in any subclone (not shown). Together, these data demonstrate that the short $\gamma 2b$ transcripts produced in the PD line do not contain variable region sequences. Previously, it was shown that the short $\gamma 2b$ transcripts produced in the A-MuLV-transformed pre-B cell lines 3-1 and 18-8 (Alt *et al.*, 1982a,b) also do not contain variable region sequences. Transcripts arising from the unrearranged α and $\gamma 3$ constant region genes were not detectable in our set of A-MuLV transformants (not shown). Thus specific transcriptional activation of the unrearranged $\gamma 2b$ locus seems to be a general property of A-MuLV-transformed pre-B cell lines; notably, of the six lines which produce the highest levels of $\gamma 2b$, the four lines analyzed for class-switching (18-8, 18-48, PD and 2M3) have all shown evidence of a specific switch to $\gamma 2b$ (Alt *et al.*, 1982b; Siden *et al.*, 1979; this report; and data not shown).

Most of the apparently novel aspects of gene rearrangement and expression discovered in A-MuLV-transformed pre-B cell lines have been observed to occur during normal B cell development (reviewed in Yancopoulos and Alt, 1986; Alt *et al.*, 1986). We examined the neonatal liver, which is a primary B cell differentiation organ rich in pre-B cells and immature B lymphocytes, for expression of $C_{\gamma 2b}$ -containing transcripts. Short $C_{\gamma 2b}$ -hybridizing transcripts, which were the same size as the germline $C_{\gamma 2b}$ transcripts seen in pre-B cell lines, were readily detectable in neonatal liver; however, $C_{\gamma 2b}$ -hybridizing transcripts the size of normal $\gamma 2b$ mRNA were not observed (Figure 4A,B). Thus, these findings support the possibility that the production of truncated $\gamma 2b$ transcripts in A-MuLV transformants may reflect a normal physiological phenomenon.

Discussion

Implications of secondary site-specific joins at the IgH locus

We examined large numbers of subclones from several independent A-MuLV transformants for secondary rearrangements of initially productive V_HDJ_H joins which resulted in altered H-chain protein expression patterns. Although we did not find the V_H -replacement events that this search was designed to elucidate, our analyses of the secondary rearrangement events observed have provided novel insights into aspects of the variable region gene assembly and H-chain class switch processes. Parallel analyses of subclones of a sister line to PD, 300-19, demonstrated the occurrence of V_H -replacement events in that line (Reth *et al.*, 1986a). Replacement of V_xJ_x rearrangements by joining of an upstream V_x to a downstream J_x (Lewis *et al.*, 1982; Feddersen and Van Ness, 1985; M.G.Reth and F.W.Alt, unpublished data), replacement of primary DJ_H rearrangements by joining

of an upstream D to a downstream J_H (Reth *et al.*, 1986b), and site-specific secondary joins of non-V_κ elements to sequences near the J_κ locus (Hochtl and Zachau, 1983; Durdik *et al.*, 1984; Siminovich *et al.*, 1985) have already been observed.

The secondary rearrangement event in PD31 has replaced a previously productive V_HDJ_H rearrangement by a site-specific join between a LINE-1 sequence and a remaining J_H segment. This event argues that regulation of H-chain allelic exclusion is not mediated by either a shut-down of the site-specific recombinase enzyme(s) or by preventing access of such enzymes to the DNA in the vicinity of the J_H locus. However, these findings are not inconsistent with the idea that prevention of additional V_H to DJ_H joins in a cell with a productive V_HDJ_H rearrangement and a remaining DJ_H intermediate is mediated by specifically denying recombinase 'access' to unrearranged V_H segments (Yancopoulos and Alt, 1985; Yancopoulos *et al.*, 1986). For example, V[H] replacement events have been demonstrated to frequently involve the most proximal unrearranged V_H segment (Reth *et al.*, 1986a), and it has been demonstrated that the activating effect of the IgH enhancer can extend to such proximal unrearranged V_H segments (Wang and Calame, 1985). Therefore, V_H segments immediately upstream to an initial V_HDJ_H rearrangement may remain accessible for such secondary rearrangement events while the rest of the V_H locus is 'inaccessible'. Thus, 'accessible' recombination recognition sequences outside of the 'inaccessible' V_H locus may occasionally be utilized in rearrangements involving the J_H locus, such as with V-replacement or the join described here, even after the formation of a productive V_HDJ_H rearrangement. Supporting the notion that the LINE-1 sequence involved in the VRS-JH3 join is accessible in pre-B cells, it is known that LINE sequences are actively transcribed in a variety of cells (Singer and Skowronski, 1985). It has also been proposed that LINE sequences may foster non-specific rearrangements in their vicinity (Singer and Skowronski, 1985), and retroviruses may preferentially integrate downstream of LINE-1 sequences (Okamoto *et al.*, 1986; Sodroski *et al.*, 1984). A potential recombination-enhancing activity of LINE repeats may also play a role in the unexpected rearrangement of a LINE sequence by the site-specific recombinase of pre-B cells.

The presence of potential recombination recognition sequences situated similarly within both mouse and human LINEs may be fortuitous or may serve an as yet undefined purpose. Site-specific joining events between J_H segments and 'pseudo-recognition sequences' near oncogenes have recently been implicated in the oncogenesis of several human B-lineage tumors (Tsujiimoto *et al.*, 1985). Our current results, coupled with our previous recombination substrates studies (Yancopoulos *et al.*, 1986), suggest that such 'aberrant' rearrangement events at the J_H locus may be quite common and are not limited to pre-B cell stages actively involved in V_H to DJ_H joining — they can even occur in cells already producing μ H-chain. Notably, the controversial Blym-1 oncogene is actually a portion of a mouse LINE-1 element whose purported transforming potential is proposed to have been activated by rearrangements imposed during molecular cloning (Rogers, 1986; Cooper *et al.*, 1986). It has been suggested that *cis*-activation of human LINE-1 transcription by downstream integration of the HTLV-1 provirus may play a role in HTLV-I leukemogenesis (Okamoto *et al.*, 1986), possibly by affecting the proposed role of LINE-1 transcripts in the coordinate transcriptional regulation of cellular genes (Davidson and Britten, 1979). The juxtaposition of LINE and IgH coding sequences could explain the large Ig-cross-reactive protein produced by PD31 as a unique fusion pro-

tein product. Such a fusion protein may be useful in identifying a heretofore undescribed protein production of LINE-1 elements, and it is intriguing to consider the possibility that isolation of cellular subclones expressing this fusion protein was aided by a selective growth advantage conferred by the fusion protein.

Kappa gene rearrangements in the absence of normal μ protein production

Production of μ protein is associated with the onset of kappa gene rearrangement in a sister cell line to PD, the 300-19 line, leading to the proposal that the μ protein provides a signal prerequisite for initiating rearrangement at the kappa locus (Reth *et al.*, 1985). The loss of normal μ production in PD31 does not prevent a high rate of continuing kappa rearrangement in this line (Lewis *et al.*, 1982). Although it is possible that production of μ protein was not prerequisite to the onset of kappa rearrangement in this line, based on our previous findings (e.g. Reth *et al.*, 1985) we consider several other likely possibilities. For example, it is possible that the putative fusion protein produced by PD31 continues to supply a necessary signal for kappa rearrangement; high levels of the short μ proteins produced from DJ_H rearrangements can apparently also supply this signal (Reth *et al.*, 1985). Alternatively, the initial signal provided by μ production in the PD parent line may result in irreversible progression to the kappa-rearranging cell stage. The example of PD31 confirms that identification of kappa gene rearrangements or κ protein production in non- μ -producing cell lines does not negate the possibility of a role for μ chain production in the onset of light chain rearrangement (see Ritchie *et al.*, 1984); thus, when a lineage analysis is possible as with PD31, many such 'light-chain-only' lines may be found to have produced μ chains at some point in their history.

Heavy chain class switching in pre-B cells

For several reasons, among them the high frequency and S-region-mediated nature of the joins, class switching in A-MuLV transformants was proposed to be more representative of the normal physiological switch event than the low frequency homologous recombination-mediated events observed in myelomas (DePinho *et al.*, 1984). The switch in the PD6 line seems to satisfy these criteria, suggesting that the complex inversion involved in the switch process may have physiological relevance; a c-myc translocation to the S _{μ} region resulted in a similar inversion (Greenberg *et al.*, 1982). Deletions within the S _{μ} region frequently occur in cell lines which undergo class switch events; the similar S _{γ 2b} deletions described here suggest that these deletions may commonly occur within 'accessible' S regions in cells with high levels of 'switch recombinases'.

Switching to γ 2b seems to be a high frequency event in pre-B cell lines — all A-MuLV transformants described so far have switched to this region by a recombination/deletion mechanism (see above). This isotype-specific switching occurs spontaneously and is clearly antigen and T cell independent. We present evidence that preferential switching to γ 2b production in pre-B cell lines is correlated with specific transcription of the germline γ 2b locus. Although we have not defined the structure of these transcripts in detail, the lack of V_H portions suggest that they may be analogous in structure to the 'sterile' μ constant region transcripts that originate in the S _{μ} region (Lennon and Perry, 1986). Sterile μ constant region transcripts also occur in all of the A-MuLV transformants described in this report (Alt *et al.*, 1982a). Such transcription may reflect a specific increase in the 'accessibility' of both the S _{μ} and S _{γ 2b} locus to switch-recombination enzymes, targeting pre-B cells to switch to γ 2b production. The short γ 2b transcripts we find in normal neonatal liver are consistent with

the possibility that the mechanisms responsible for preferential switching to $\gamma 2b$ in pre-B cell lines may be operative *in vivo*, although the identity of the sequences observed in normal tissues and the A-MuLV transformants remains to be confirmed. Because normal-sized $\gamma 2b$ RNA was not detected in the neonatal liver, the short $\gamma 2b$ transcripts may reflect a pre-commitment of descendants of the early stage B cells in the neonate to switch to $\gamma 2b$; the actual switch event may require an additional event, such as the expression of a common switch recombinase.

Germline $C_{\gamma 2b}$ transcription raises the intriguing possibility that a transcriptional (and recombinational) enhancer element active in pre-B cells is located near the $C_{\gamma 2b}$ locus. Similar elements, under the control of different *trans*-acting factors, may be associated with other C_H regions and may be involved in directing isotype-specific switch events. For example, it has recently been proposed that a specific increase in the accessibility of the S_{α} and S_{ϵ} regions to switch recombinases targets a B-cell lymphoma to preferentially switch to IgA and IgE production (Stavnezer-Nordgren and Sirlin, 1986). Differential activation of C_H -associated enhancer elements may allow a single switch recombinase to differentially perform isotype-specific switches, obviating the need for sets of class-specific switch recombinases to explain targeted switching to specific H-chain classes. Thus the distinct recombinational systems in lymphocytes which perform variable region gene assembly and H-chain class switching may both be directed by the 'accessibility' of their substrate gene segments.

Materials and methods

Cells

Details concerning the derivation, growth and characteristics of these lines have been presented elsewhere (Alt *et al.*, 1984; Lewis *et al.*, 1982). The adult spleen and neonatal liver samples described were prepared from pools of these tissues isolated from C57BL/6J mice in all cases. Four litters were pooled for each neonatal liver sample; timed pregnant mice were used to obtain these litters. Four adult spleens were pooled for each adult spleen sample. C57BL/6J mice were obtained from the Jackson Laboratory, Bar Harbor, Maine.

Protein analysis

Analysis of Ig protein production by these lines was performed as previously described (Lewis *et al.*, 1982).

Analysis of heavy chain gene organization

DNA and RNA preparation, restriction enzyme digestions, agarose gel electrophoresis, DNA and RNA blotting procedures, probe preparations and hybridization conditions were performed as previously described (Yancopoulos *et al.*, 1984; Yancopoulos and Alt, 1985). The J_H probe used was previously described (Alt *et al.*, 1984) as was the 3'- $S_{\gamma 2b}$ probe (DePinho *et al.*, 1984) and the $C_{\gamma 2b}$ probe (Bothwell *et al.*, 1981). The V_A probe specific for the expressed variable region in PD was prepared by isolating the 5' *NcoI/BamHI* fragment from the cloned A rearrangement (see Figure 2B).

Genomic cloning and DNA sequencing

Genomic libraries were prepared by cloning complete *EcoRI* digests of genomic DNA from the various lines into the $\chi 16A$ vector as was done previously (Yancopoulos *et al.*, 1984); screening with appropriate probes, plaque purification, isolation and subcloning of inserts and Maxam and Gilbert sequencing (Maxam and Gilbert, 1977) of relevant portions of these inserts were performed as previously described (Yancopoulos *et al.*, 1984).

DNA sequence homology search

A computer search for all murine sequences in GenBank (provided by the National Institutes of Health) homologous to the VRS was performed using the Microgenic Sequence Analysis Program (copyrighted in 1983 by SciSoft, Inc.).

Acknowledgements

We thank Warren Pear and George Klein for careful reading of this manuscript and for initially suggesting that the VRS repetitive element was similar to the LINE repeats. This work was supported by a Searle Scholars Award and N.I.H. grants AI-20047, CA-40427 and CA-2112-09 to F.W.A., who is an Irma T. Hirsch Career Scientist and a Malinckrodt Scholar. N.R. was supported by N.I.H. grants CA-24530 and CA-33771, R.A.D. was supported by N.I.H.

Physician-Scientist award AI-00602-03, K.A.Z. was supported by N.I.H. training grant CA 09503, and S.G.L. and G.D.Y. were supported by N.I.H. training grant GM-07367.

References

- Akira, S., Sugiyama, H., Yoshida, M., Kitkutani, H., Yamamura, Y. and Kishimoto, T. (1983) *Cell*, **34**, 545–556.
- Alt, F., Rosenberg, N., Lewis, S., Thomas, E. and Baltimore, D. (1981) *Cell*, **27**, 381–390.
- Alt, F.W., Rosenberg, N., Enea, V., Siden, E. and Baltimore, D. (1982a) *Mol. Cell. Biol.*, **2**, 386–390.
- Alt, F.W., Rosenberg, N., Casanova, R.J., Thomas, E. and Baltimore, D. (1982b) *Nature*, **296**, 325–331.
- Alt, F.W., Yancopoulos, G.D., Blackwell, T.K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S. and Baltimore, D. (1984) *EMBO J.*, **3**, 1209–1219.
- Alt, F.W., Blackwell, T.K., DePinho, R.A., Reth, M.G. and Yancopoulos, G.D. (1986) *Immunol. Rev.*, **89**, 5–30.
- Bernard, O., Hozumi, N. and Tonegawa, S. (1978) *Cell*, **15**, 1133–1144.
- Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. (1981) *Cell*, **24**, 625–637.
- Burrows, P.D., Beck-Engeser, G.B. and Walb, W.R. (1983) *Nature*, **306**, 243.
- Cooper, G.M., Goubin, G., Diamond, A. and Neiman, P. (1986) *Nature*, **320**, 579–580.
- Davidson, E.H. and Britten, R.J. (1979) *Science*, **204**, 1052–1059.
- DePinho, R., Kruger, K., Andrews, N., Lutzker, S., Baltimore, D. and Alt, F.W. (1984) *Mol. Cell. Biol.*, **4**, 2905–2910.
- Durdik, J., Moore, M.W. and Selsing, E. (1984) *Nature*, **307**, 749–752.
- Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980) *Cell*, **19**, 981–992.
- Fedderson, R.M. and Van Ness, B.G. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4793–4797.
- Gebhard, W., Meitinger, T., Hochtl, J. and Zachau, H.G. (1972) *J. Mol. Biol.*, **157**, 453–471.
- Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) *Cell*, **33**, 717–728.
- Greenburg, R., Lang, R.B., Diamond, M.S. and Marcu, K.B. (1982) *Nucleic Acids Res.*, **10**, 7751–7761.
- Hattori, M., Kuhara, S., Takenaka, O. and Sakaki, Y. (1986) *Nature*, **321**, 625–628.
- Hochtl, J. and Zachau, H.G. (1983) *Nature*, **302**, 260–263.
- Kataoka, T., Miyata, T. and Honjo, T. (1981) *Cell*, **23**, 357–368.
- Lennon, G.G. and Perry, R.P. (1985) *Nature*, **318**, 475–477.
- Lewis, S., Rosenberg, M., Alt, F. and Baltimore, D. (1982) *Cell*, **30**, 807–816.
- Loeb, D., Padgett, R.W., Hardies, S.C., Shehee, W.R., Comer, M.B., Edgell, M.H. and Hutchinson, C.A. (1986) *Mol. Cell. Biol.*, **6**, 168–182.
- Manuelidis, L. (1982) *Nucleic Acids Res.*, **10**, 3211–3219.
- Marcu, K. and Cooper, M. (1982) *Nature*, **306**, 243.
- Maxam, A. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 1800–1804.
- Okamoto, T., Reitz, M.S., Clarke, M.F., Jagodzinski, L.L. and Wong-Staal, F. (1986) *J. Biol. Chem.*, **261**, 4615–4619.
- Pernis, B.G., Chiappino, G., Kelus, A.S. and Gell, P.G.H. (1965) *J. Exp. Med.*, **122**, 853–876.
- Rathbun, G.A. and Tucker, P.W. (1986) In Kelsø, G. and Schultz, D. (eds), *Evolution and Vertebrate Immunity: The Antigen Receptor and MHC Gene Families*. Academic Press, in press.
- Reth, M.G., Amirati, P., Jackson, S. and Alt, F.W. (1985) *Nature*, **317**, 353–355.
- Reth, M., Gehrmann, P., Petrac, E. and Wiese, P. (1986a) *Nature*, **322**, 840–842.
- Reth, M.G., Jackson, S. and Alt, F.W. (1986b) *EMBO J.*, **5**, in press.
- Ritchie, K.A., Brinster, R.L. and Storb, U.L. (1984) *Nature*, **312**, 517–520.
- Rogers, J. (1986) *Nature*, **320**, 579.
- Rusconi, S. and Kohler, G. (1985) *Nature*, **314**, 330–334.
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. and Tonegawa, S. (1980) *Nature*, **286**, 676–683.
- Shimizu, A. and Honjo, T. (1984) *Cell*, **36**, 801–803.
- Shimizu, A., Takahashi, N., Yaiota, Y. and Honjo, T. (1982) *Cell*, **24**, 499–506.
- Siden, E.J., Baltimore, D., Clark, D. and Rosenberg, N.E. (1979) *Cell*, **16**, 389–396.
- Siminovitch, K.A., Bakhski, A., Goldman, P. and Korsmeyer, S.J. (1985) *Nature*, **316**, 260–261.
- Singer, M.F. and Skowronski, J. (1985) *Trends Biochem. Sci.*, **10**, 119–121.
- Stavnezer-Nordgren, J. and Sirlin, S. (1986) *EMBO J.*, **5**, 95–102.
- Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E. and Croce, C.M. (1985) *Science*, **229**, 1390–1393.
- Wang, X.F. and Calame, K. (1985) *Cell*, **43**, 659–665.
- Weaver, D., Costantini, F., Imanishi-Kari, T. and Baltimore, D. (1985) *Cell*, **42**, 117–127.

- Wu, T.T., Reid-Miller, M., Perry, H.M. and Kabat, E.A. (1984) *EMBO J.*, **3**, 2033–2040.
- Yancopoulos, G.D. and Alt, F.W. (1985) *Cell*, **40**, 271–281.
- Yancopoulos, G.D. and Alt, F.W. (1986) *Annu. Rev. Immunol.*, **4**, 339–368.
- Yancopoulos, G.D., Desiderio, S.V., Paskind, M., Kearney, J.F., Baltimore, D. and Alt, F.W. (1984) *Nature*, **311**, 727–733.
- Yancopoulos, G.D., Blackwell, T.K., Suh, H., Hood, L.E. and Alt, F.W. (1986) *Cell*, **44**, 251–259.

Received on 9 April 1986; revised on 19 August 1986