Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 2689–2693, April 1986
Medical Sciences

A DNA insertion/deletion necessitates an aberrant RNA splice accounting for a μ heavy chain disease protein

(immunoglobulin genes)

AJAY BAKHSHI*, PAUL GUGLIELMI*†, ULI SIEBENLIST‡, JEFFREY V. RAVETCH§, JANE P. JENSEN*, AND STANLEY J. KORSMEYER*

*Metabolism Branch, National Cancer Institute, ‡Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; and §DeWitt Wallace Research Laboratory, Memorial Sloan Kettering Cancer Center, New York, NY 10021

Communicated by Matthew D. Scharff, December 18, 1985

ABSTRACT **The human heavy chain disease protein BW is an immunoglobulin μ-chain variant whose amino terminus is initiated at the fifth amino acid of the first constant region domain. We cloned and analyzed both rearranged heavy chain alleles from BW leukemic cells to determine the molecular basis for this deleted protein. The phenotypically excluded heavy-chain allele possessed two intermediate recombinations of separate variable-diversity (V-D) and diversity-joining (D-J) junctions, neither of which were expressed. The productive allele, responsible for the μ chain, had a complete V-D-J_4 recombination but as a result of a single-base deletion possessed stop codons within the variable region. More important, a small DNA insertion/deletion eliminated the J_4 donor splice site. This necessitated an aberrant RNA splice between the leader region and the first constant region domain creating a shortened 2.35-kilobase μRNA. A recognition sequence for signal peptidase predicted a cleavage at the fifth amino acid of the first constant region domain. These molecular events are responsible for the truncated μ chain that lacks a variable region and fails to assemble light chains.**

The human heavy chain diseases (HCDs) are proliferative disorders of B-lymphoid cells that produce truncated monoclonal immunoglobulin heavy (H) chain molecules (1, 2). HCD proteins of the 3 major immunoglobulin classes: γ, μ, and α have been described. The abnormal H chains in μHCD have had deletions involving one, two, or three of the amino-terminal protein domains (3). The analysis of variant molecules has often yielded insights into the processes responsible for the synthesis and regulation of normal proteins. In the case of variant immunoglobulin heavy chains prior work has discounted extensive postsynthetic degradation (1, 2). To express a complete H chain the separate gene subsegments of the variable (V_H), diversity (D_H), and joining (J_H) regions must be correctly recombined at the DNA level, and the remaining introns need to be removed by RNA splicing (4). Thus the abnormal H chains could be the result of either aberrant gene rearrangements or of mutational events leading to altered RNA splicing or incorrect transcription initiation. To investigate these alternatives we undertook a molecular analysis of the HCD protein from patient BW with μHCD.

Structural analysis of the μHCD protein present in the serum of the patient BW had revealed that the monomeric H chain, with a molecular weight of 58,000, lacked associated (light) L chains. The entire variable region was deleted, with the amino-terminal sequence being initiated at the fifth amino acid of the first constant region domain (C_{H1}) of the constant region of the μ chain (C_μ). Furthermore, significant postsynthetic degradation was not occurring because the primary translation product, the cytoplasmic and the secreted forms of the μ chain were similarly truncated (5). To determine the molecular basis for this μHCD, we cloned and analyzed both rearranged H chain alleles from BW leukemic cells. We demonstrate that an unexpected DNA insertion/deletion event forced an abnormal RNA splice resulting in a truncated μ chain.

MATERIALS AND METHODS

Isolation and Analysis of Nucleic Acids. High molecular weight DNA was isolated from nuclei of BW leukemic cells, digested with the appropriate restriction endonuclease, size-fractionated by agarose gel electrophoresis, and transferred to nitrocellulose sheets (6). Blots were hybridized with nick-translated or random-primed ^{32}P-labeled DNA fragments and washed as described (7, 8).

Total cellular RNA was prepared from the leukemic cells by resuspending the cells in 4 M guanidine thiocyanate, disrupting them with a Polytron homogenizer (Brinkmann), and centrifuging the RNA through a 5.7 M $CsCl_2$ gradient (9). RNA was electrophoresed in 7% (vol/vol) formaldehyde/1% agarose gels, blotted to nitrocellulose filters, and hybridized to ^{32}P-labeled DNA probes as described (10).

DNA Probes. The C_μ probes utilized included 1.3-kilobase (kb) *Eco*RI and 0.9-kb *Eco*RI genomic fragments encompassing C_{H1}-C_{H2}-C_{H3} and C_{H3}-C_{H4} coding regions, respectively (11). The μ membrane probe was a 0.8-kb *Sac* I–*Bam*HI genomic fragment that contained both exons for the membrane anchoring piece (11). The J_H probe consisted of a 2.3-kb germ-line *Sau*3A genomic fragment, and the constant region of the κ chain gene probe was a 2.5-kb *Eco*RI genomic fragment (8). The D_H segment probe was a 1.8-kb germ-line *Bam*HI fragment (12).

Isolation of the Rearranged Immunoglobulin Alleles. DNA from the leukemic B cells of patient BW was partially digested with *Mbo* I and size-fractionated by centrifugation through a 10–40% (wt/vol) sucrose gradient. The gradient fractions containing DNA fragments of 10–15 kb were ligated to the *Bam*HI arms of the phage vector Charon 28. The packaging of phage DNA and the subsequent plating and screening of plaques by hybridization were all according to standard procedures (13). Both rearranged alleles were isolated and further characterized by restriction mapping and DNA sequencing.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HCD, heavy chain disease; C_H, constant region domain of the heavy chain; C_μ, constant μ region; H chain, heavy chain; L chain, light chain; CLL, chronic lymphocytic leukemia; V_H, D_H, and J_H, the variable, diversity, and joining segments of the H chain genes; kb, kilobase(s); bp, base pair(s).

†Present address: Laboratory of Immunochemistry and Immunopathology, Hospital Saint-Louis, 75475, Paris, Cedex 10, France.

DNA Sequencing. DNA fragments were ligated into M13 phage and transformed into competent cells, and relevant clones were identified by Benton–Davis hybridization, using a J_H probe (14). Sequencing was performed using the dideoxy chain termination method (15). In addition, some regions were sequenced using the chemical degradation method of Maxam and Gilbert (16).

RESULTS

The RNA Coding for the HCD Protein BW Is Truncated. RNA extracted from BW leukemic cells and incubated with a C_μ probe revealed a predominant 2.35-kb message (Fig. 1). μHCD is a variant of chronic lymphocytic leukemia (CLL), and CLL cells normally produce predominantly the 2.7-kb membrane form of μRNA (μ_mRNA) as seen in patient DC with CLL (DC-CLL) (17, 18). This suggested that the final processed RNA in BW had a deletion of approximately 350 base pairs (bp). However, cells that secrete immunoglobulin also have a 2.4-kb secreted form of μRNA (μ_sRNA), as seen in the lymphoblastoid cell line RPMI 8392 (18). To demonstrate that the 2.35-kb message seen in BW represented a truncated μRNA and not simply a secreted form of message, we utilized a *Sac* I–*Bam*HI fragment that contained both exons coding for the membrane anchoring piece as a probe. This probe recognized a 2.7-kb μ_mRNA in DC-CLL and RPMI 8392 cells and a 2.35-kb μ_mRNA in BW cells. Thus the μRNA in the BW cells was truncated. Furthermore, the precursor forms of μRNA were apparently normal in size as visualized on longer exposures (data not shown). This suggested that the deletion occurred at the level of RNA processing and not earlier, during transcription. In contrast, the κRNA and its precursor forms were normal in size (Fig. 1). This κRNA directed the synthesis of a normal-sized κ protein that was detected in the urine of BW as Bence–Jones protein (5).

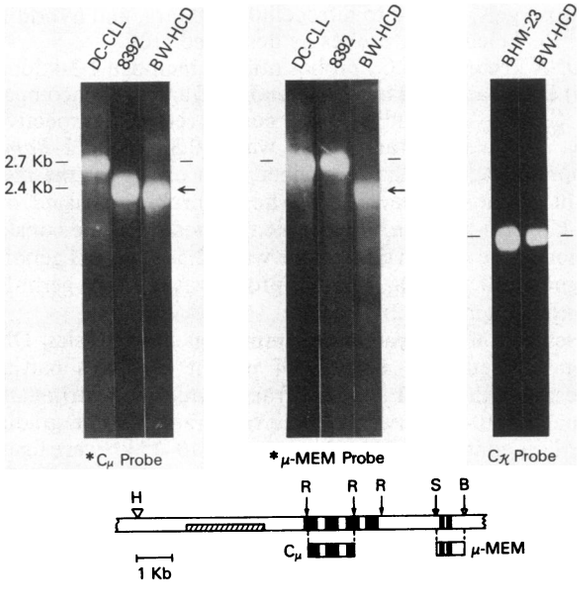

FIG. 1. Blot–hybridization analysis of μHCD RNA from BW cells. Total RNA (10 μg) from DC-CLL, a B-cell line (RPMI 8392), and BW leukemic cells was examined using the 1.3-kb C_μ genomic fragment (*Left*) and the 0.8-kb μ-membrane (μ-MEM) fragment (*Center*) shown in the schematic. A 2.5-kb genomic C_κ fragment was used in the analysis of BHM-23 (a $\mu\kappa$ producing cell line) and BW (*Right*). S, *Sac* I; R, *Eco*RI; B, *Bam*HI; H, *Hin*dIII. A 2.7-kb μ_mmRNA species is seen in RPMI 8392 with the C_μ probe on longer exposures.

Constant μ Domains Were Intact. All four C_μ region domains were intact and were expressed in BW μRNA when analyzed by a combination of Southern blotting, S1 nuclease analysis, and DNA sequencing. When DNA from BW leukemic cells was digested with the restriction endonuclease *Sac* I and probed with a C_μ probe both BW and the control revealed a single 4.5-kb band (Fig. 2*a*). This 4.5-kb fragment encompasses all four C_μ domains. Further dissection with *Eco*RI revealed that a 1.3-kb fragment containing C_{H1}, C_{H2}, most of C_{H3}, and a 0.9-kb fragment containing the remainder of C_{H3} and all of C_{H4}, were also retained in their germ-line form in BW DNA (Fig. 2*a*). Any exon deletion or internal

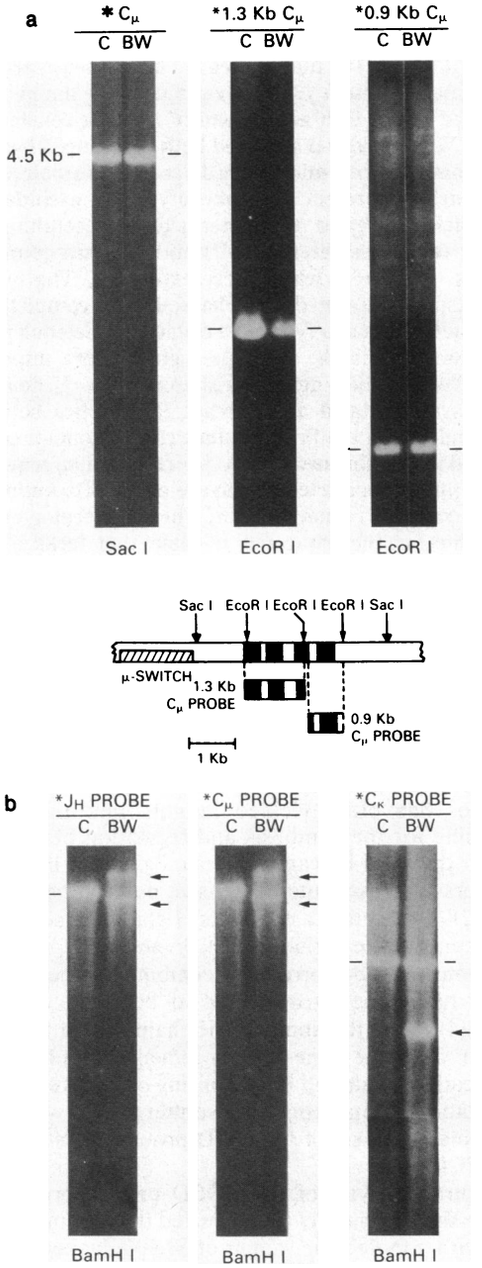

FIG. 2. Southern blot analysis of μHCD DNA from BW cells. (*a*) Leukemic B-cell DNA was digested with the restriction enzymes indicated and the filters were hybridized with radiolabeled 1.3-kb or 0.9-kb C_μ fragments shown in the schematic. C, germ-line control. (*b*) Rearrangements of both H-chain alleles were detected with J_H and C_μ probes (arrows) as compared to the germ-line genes (indicated by tick marks). With the C_κ probe single κ chain gene rearrangement was present (arrow). C, germ-line control.

rearrangement would have generated a new band. However, since BW retained these fragments in their germ line form no significant constant region DNA deletion had occurred. Furthermore, S1 nuclease analysis and DNA sequence analysis of the expressed genomic C_μ allele (data not shown) confirmed that all four C_μ domains were intact, had proper donor and acceptor splice signals, and were expressed in BW mRNA.

Organization of the Rearranged H-Chain Alleles. Since the C_μ region was intact and the BW protein lacked variable region information, we turned our attention to the gene subsegments coding for the variable portion of the H chain. When DNA from the leukemic cells was digested with *Bam*HI and hybridized to a J_H probe, both alleles were rearranged (Fig. 2*b*). Such rearrangements represent attempts at assembling the gene subsegments of the variable portion of the immunoglobulin molecule. In addition, both rearranged J_H alleles also possessed C_μ information (Fig. 2*b*). Thus, at this level of analysis the rearranged allele responsible for the abnormal RNA and protein could not be distinguished. In contrast, only one κ gene allele was rearranged and was thus responsible for the normal-sized κRNA

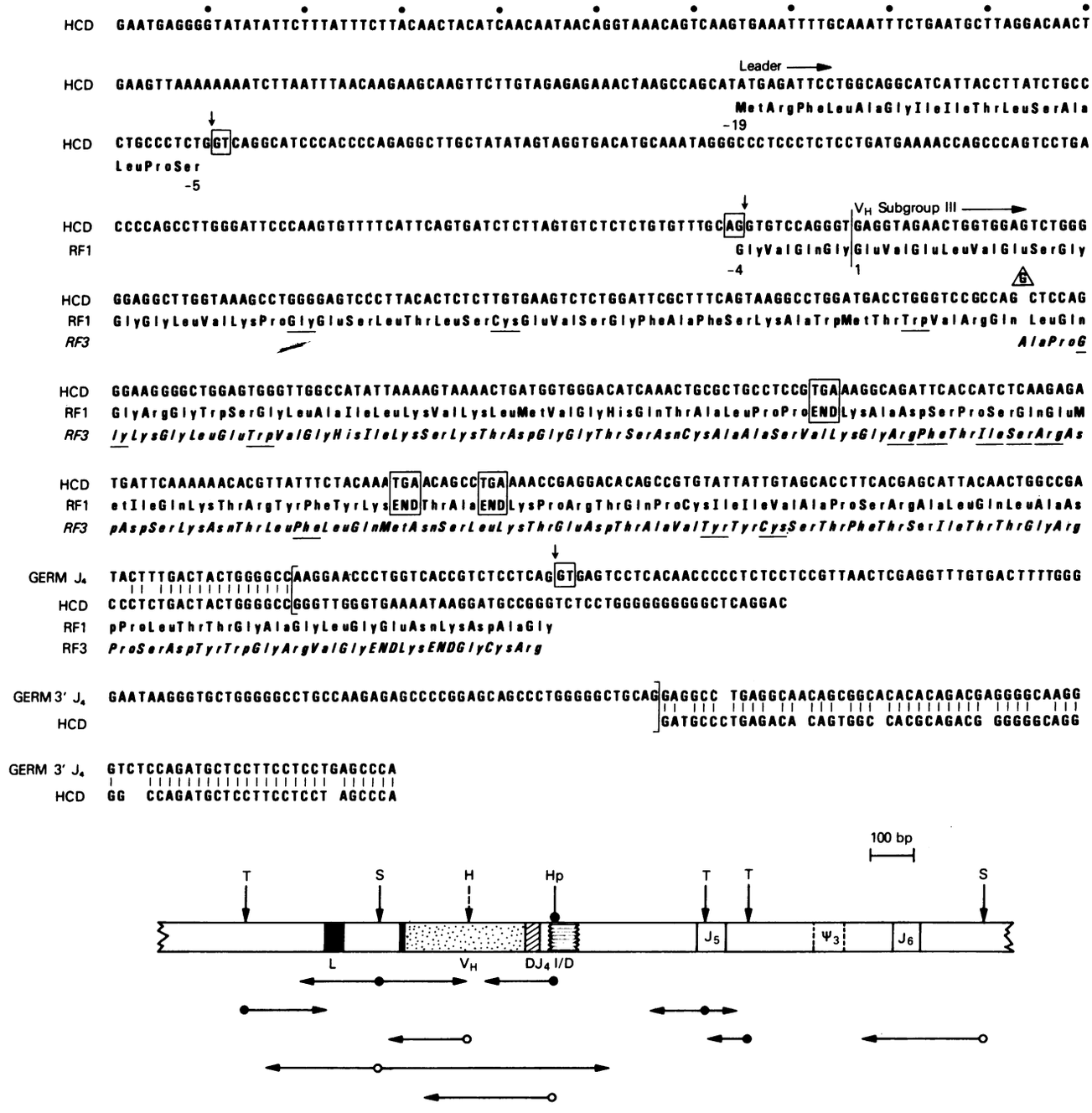

FIG. 3. (*Upper*) Nucleotide sequence of the V-D-J_4 recombination region. The deduced amino acid sequence in two reading frames, reading frames 1 and 3, is shown. The site of the single-base deletion (ΔG) is indicated. The three stop codons in reading frame 1 are boxed, and the putative RNA donor and acceptor splice sites are indicated by the vertical arrows. The DNA insertion/deletion within the J_4 coding region is demarcated by brackets. (*Lower*) The restriction map and sequencing strategy of the V-D-J_4 recombination. Open circles, sequencing using M13 dideoxy chain termination; closed circles, restriction sites that were 3′ or 5′ end-labeled and sequenced according to Maxam and Gilbert. S, *Sau*3A; T, *Taq* I; Hp, *Hpa* I; H, *Hae* III. The invariant amino acids for a subgroup III V_H region are underlined in reading frame 1 and 3, respectively.

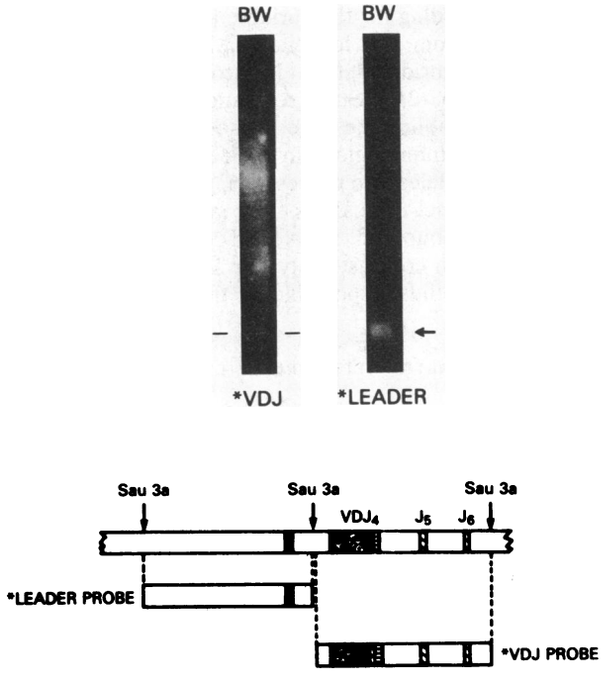

FIG. 4. Expression of the V-D-J_4 recombination. Total RNA (10 μg) per lane was used, and the blots were hybridized with radiolabeled *Sau*3A genomic fragments containing either the V-D-J_4 recombination (*Left*) or the leader region (*Right*). Under the blots is a map showing the probes used.

produced (Fig. 2*b*). To directly analyze the rearranged H-chain alleles a genomic library of the leukemic cell DNA was constructed and both alleles were isolated by replica screening 600,000 recombinant clones with radiolabeled J_H and C_μ probes. The organization and restriction map of the H-chain allele responsible for the μ chain production is shown in Fig. 3. This allele had a complete V-D-J recombination. The other allele had intermediate recombinations with separate V-D and D-J joinings (U.S., unpublished results). Neither of these incomplete recombinations hybridized to the short BW μRNA and thus were not responsible for the μHCD protein.

Single-Base Deletion in V-D-J Generates Stop Codons. The nucleotide sequence of the V-D-J recombination is presented in Fig. 3. The deduced amino acid sequence reveals a protein consistent with a subgroup III variable region (19). The invariant residues, glycine-15 and cysteine-22, of the framework region 1 of a subgroup III V_H were present in reading frame 1. The framework region 2 started appropriately with the invariant tryptophan-36 and the glycine-42, but the predicted invariant tryptophan-47 was replaced by a glycine. Furthermore, the invariant residues in framework region 3 were altered, and three translation stop codons were present in this reading frame as a result of a single-base deletion (Fig. 3). If the third reading frame is examined from the site of this deletion the invariant residues as per the numbering convention of Kabat (glycine-42, tryptophan-47, arginine-66, phenylalanine-67, isoleucine-69, serine-70, arginine-71, tyrosine-90, and cysteine-92) would all be present (Fig. 3) (19). The invariant tyrosine-79 is replaced by phenylalanine, perhaps as a result of a single adenosine to thymidine mutation. Stop codons do exist in any of the three potential reading frames. Thus a single-base deletion appears to shift the usual reading frame within this rearranged V_H gene and generate stop codons.

An Insertion/Deletion Alters the J_4 Donor Splice Site. The V_H region had recombined site specifically with a D_H region and the J_4 gene subsegment. The J_5 and J_6 regions were unaltered in their germ-line position. The 5′ portion of the joining region within the recombined V-D-J revealed 16 of 19 bp that were homologous to the sequence of the germ-line J_4 region (Fig. 3). However, abruptly after the sequence GG-GCC, there was no homology between the joining segment in the recombined V-D-J_4 and the normal J_4. Moreover, there was no homology between a 51-bp region of the rearranged gene and the expected 138-bp sequence that included a portion of J_4 and its 3′ flank (11, 20). Further downstream a region 3′ of the V-D-J_4 recombination and the germ-line J_4–J_5 intron were once again homologous although differences were present that might indicate somatic mutation or polymorphic variation between individuals in a noncoding region (Fig. 3). Thus, a recombinational event had deleted 138 nucleotides and inserted 51 bp, of unknown origin, in their place. This insertion/deletion altered the translation in all reading frames, such that the invariant serine-112 was not present (19). More importantly, the J_4 donor splice sequences had been altered from CAGGTGAGT to GGGTCTCCT, which would be nonfunctional (21). Thus this V-D-J_4 would not be properly spliced to the C_{H1} exon to give a final processed C_μ message. Correspondingly, when this V-D-J segment was used as a probe against RNA from BW cells, it failed to hybridize (Fig. 4). Conversely, a 5′ probe encompassing the

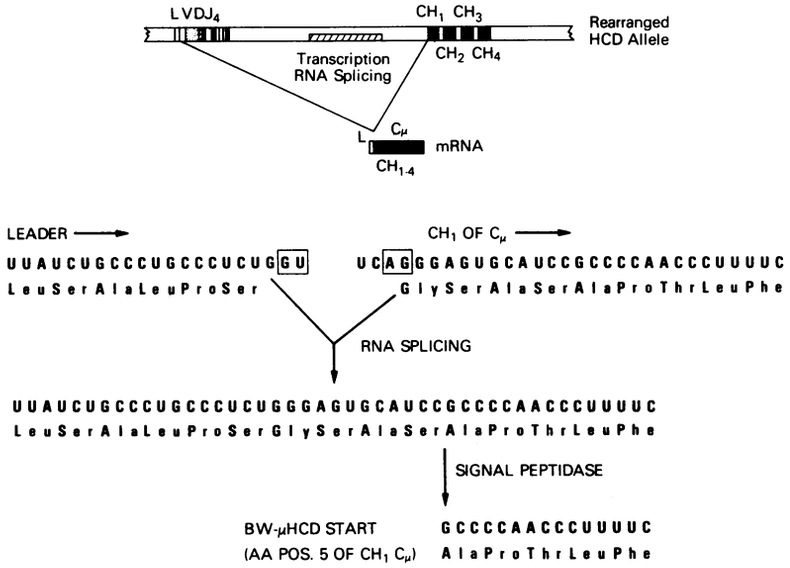

FIG. 5. (*Upper*) Schematic representation of the aberrant RNA splice between the leader region and (C_{H1}) of C_μ. (*Lower*) Nucleotide and amino acid sequence resulting from the alternate RNA splice. The predicted signal peptidase cleavage site is indicated, generating the μHCD protein from patient BW.

leader region did hybridize to the same small message bearing the C_μ region (Fig. 4). The donor splice site of the leader region appears to be alternatively spliced to the acceptor splice site of C_{H1}, thus eliminating the variable region coding information.

DISCUSSION

The human HCDs might predictably result from a variety of molecular errors in the complex processes of immunoglobulin gene assembly, transcription, and RNA processing. Genomic deletions that arise from errant V-D-J joining attempts or abnormal H-chain class switches could result in the loss of information for protein domains. Alternatively, somatic mutations could eliminate crucial donor or acceptor splice sites and result in these truncated proteins that are missing exon information. Our analysis of the basis for μHCD from BW cells revealed multiple molecular defects. The dominant mechanism creating the variant H chain was a DNA insertion/deletion that eliminated the J_4 donor splice site and necessitated an alternative splice that juxtaposed the leader segment and C_μ. In addition a single-base deletion as well as point mutations were present and resulted in the creation of stop codons within the variable region.

Elimination of the proper splice between V-D-J_4 and C_{H1} necessitated an alternate choice. The normal splice between the leader and its V_H might occur at a low frequency but would result in a large, incompletely processed RNA containing stop codons. Instead, the alternative choice predominates in which a splice between the donor site of the leader and the acceptor site of C_{H1} creates the transcript responsible for the translated product (Fig. 5). A consensus sequence for signal peptidase recognition (Gly-Ser-Gly-Ser-Ala-Ser) exists at this juncture and would result in cleavage after the last serine residue (Fig. 5) (22, 23). In fact, BW μ chain began at the fifth amino acid of C_{H1} confirming this predicted cleavage and further indicating that absolutely no proteolysis had occurred.

The J_4 donor splice site was replaced by a 51-bp inserted sequence that was not of obvious immunoglobulin origin. Moreover, while this inserted sequence was G+C rich, it bore no homology to described transposable elements (24). Furthermore, no suggestion of direct repeats, inverted repeats, or recombination signals was found at its borders (25). Insertion of repetitive elements related to intracisternal A particles has been noted in L chain genes (26). Similar elements have also been observed to activate c-*onc* genes (27).

Numerous variant H and L chains have been described and the molecular defects have been detailed in some. The mouse γ_1 H-chain mutant IF2 resulted from a genomic DNA deletion that removed the C_{H1} exon resulting in the splicing of V_H to the hinge exon (28). Similarly, the mouse α H-chain mutants Δ15 and Δ37 have DNA deletions that are felt to arise by a recombination/excision mechanism (29). Internal deletions of C_H exons have even been noted in the shortened μ chains of mutant mouse hybridomas and in the C_γ2b gene in a mutant of MPC11 (30, 31). In addition, examination of a cDNA sequence from a human γHCD, OMM, is compatible with a gene deletion model (32). Furthermore, our analysis of the genomic gene for another γHCD, RIV, revealed a deletion of most of the C_{H1} exon (33). Quite similar to μHCD from BW, the truncated κ chain of MPC11 arose from an aberrant splice of a leader to C_κ. This splice resulted because the V_κ eliminated the J_κ region by recombining at a heptanucleotide signal within the J_κ–C_κ intron (34, 35). As a consequence, the entire variable region exon is deleted from the final processed RNA. Our results indicate that a small insertion/deletion event in μHCD, BW eliminates the donor splice site at the 3′ end of a V-D-J rearrangement and necessitates an alternate RNA splice between the leader and C_{H1}. This molecular defect generates a truncated μ chain present both as a transmembrane and as a secreted molecule that lacks the variable region and fails to associate with L chain.

We wish to thank Dr. Arnold W. Strauss, Washington University, for valuable discussions and Janet Shelton for her assistance in the preparation of this manuscript.

1. Franklin, E. C. & Frangione, B. (1975) in *Contemporary Topics in Molecular Immunology*, eds. Inman, F. P. & Mandy, W. (Plenum, New York), pp. 89–126.
2. Seligmann, M., Mihaesco, E., Preud'homme, J.-L., Danon, F. & Brout, J.-C. (1979) *Immunol. Rev.* **48,** 145–167.
3. Franklin, E. C., Frangione, B. & Prelli, F. (1976) *J. Immunol.* **116,** 1194–1195.
4. Early, P., Huang, M., Davis, M., Calame, K. & Hood, L. (1980) *Cell* **19,** 981–992.
5. Bakhshi, A., Guglielmi, P., Coligan, J. E., Gamza, F., Waldmann, T. A. & Korsmeyer, S. J. (1985) *Mol. Immunol.* , in press.
6. Southern, E. M. (1979) *Methods Enzymol.* **68,** 152–176.
7. Bakhshi, A., Minowada, J., Arnold, A., Cossman, J., Jensen, J. P., Whang-Peng, J., Waldmann, T. A. & Korsmeyer, S. J. (1983) *N. Engl. J. Med.* **309,** 826–831.
8. Korsmeyer, S. J., Arnold, A., Bakhshi, A., Ravetch, J. V., Siebenlist U., Hieter, P. A., Sharrow, S. O., LeBien, T. W., Kersey, J. H., Poplack, D. G., Leder, P. & Waldmann, T. A. (1983) *J. Clin. Invest.* **71,** 301–313.
9. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rudder, W. J. (1979) *Biochemistry* **18,** 5294–5299.
10. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
11. Ravetch, J. V., Siebenlist, U., Korsmeyer, S., Waldmann, T. & Leder, P. (1981) *Cell* **27,** 583–591.
12. Siebenlist, U., Ravetch, J. V., Korsmeyer, S. J., Waldmann, T. A. & Leder, P. (1981) *Nature (London)* **294,** 631–635.
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1984) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
14. Benton, W. A. & Davis, R. W. (1977) *Science* **196,** 180–182.
15. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94,** 414–418.
16. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
17. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20,** 313–319.
18. Cossman, J., Neckers, L. M., Braziel, R. M., Trepel, J. B., Korsmeyer, S. J. & Bakhshi, A. (1984) *J. Clin. Invest.* **73,** 587–592.
19. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1983) *Sequences of Immunoglobulin Chains* (U.S. Department of Health, Education & Welfare, Washington, DC).
20. Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E. & Croce, C. M. (1985) *Science* **229,** 1390–1393.
21. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
22. Strauss, A. W. & Boime, I. (1982) *CRC Crit. Rev. Biochem.* **12** (3), 205–235.
23. Wu, H. C. (1985) in *Compartmentalization of Proteins*, eds. Strauss, A. W., Kreil, G. & Boime, I. (Springer, New York), in press.
24. Jelinek, W. R. & Schmid, C. W. (1982) *Annu. Rev. Biochem.* **51,** 813–844.
25. Weinert, T. A., Schaus, N. A. & Grindley, N. D. F. (1983) *Science* **222,** 755–765.
26. Kuff, E. L., Feenstra, A., Lenders, K., Smith, L., Hawley, R., Hozumi, N. & Shulman, M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1992–1996.
27. Canaani, E., Dreazen, O., Klar, A., Rechavi, G., Ram, D., Cohen, J. B. & Givol, D. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 7118–7122.
28. Dunnick, W., Rabbits, T. H. & Milstein, C. (1980) *Nature (London)* **286,** 669–675.
29. Dackowski, W. & Morrison, S. L. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 7091–7095.
30. Kohler, G., Potash, M. J., Lehrach, H. & Shulman, M. J. (1982) *EMBO J.* **1,** 555–563.
31. Brandt, C. R., Morrison, S. L., Birshtein, B. K. & Milcarek, C. (1984) *Mol. Cell. Biol.* **4,** 1270–1277.
32. Alexander, A., Steinmetz, M., Barritault, D., Frangione, B., Franklin, E. C., Hood, L. & Buxbaum, J. N. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 3260–3264.
33. Guglielmi, P., Bakhshi, A., Miahesco, E., Brouet, J., Waldmann, T. A. & Korsmeyer, S. J. (1984) *Clin. Res.* **32,** 348A (abstr.).
34. Choi, E., Kuehl, M. & Wall, R. (1980) *Nature (London)* **286,** 776–779.
35. Seidman, J. G. & Leder, P. (1980) *Nature (London)* **286,** 779–783.