Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3630–3634, June 1980
Immunology

Nucleotide sequence of immunoglobulin heavy chain joining segments between translocated V_H and μ constant region genes

(μ mRNA/Cμ gene/active V_H gene/J_H sequences/antibody diversity)

ORA BERNARD AND NICHOLAS M. GOUGH

Molecular Biology Laboratory, The Walter and Eliza Hall Institute of Medical Research, Post Office, Royal Melbourne Hospital, Victoria, 3050, Australia

Communicated by Sir G. J. V. Nossal, March 24, 1980

ABSTRACT **To investigate the mechanism of recombination of immunoglobulin heavy chain variable and constant region genes, we have determined the nucleotide sequence of a large portion of the recombination region between an active Cμ gene and its associated V_H gene, isolated from an IgM-secreting mouse plasmacytoma, HPC76. By comparison with the sequence of the μ mRNA, we determined the exact boundaries of the intervening sequence between the V_{H76} and Cμ genes. The rearranged V_{H76} gene encodes up to amino acid 116 without interruption, the 3′ 39 nucleotides (the J_{H76} region) being derived from an embryonic J_H segment (J_{H315}) whose sequence was recently determined [Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. (1980) *Cell* 195, 981–992]. The active J_{H76} does not use the first two codons of the embryonic J_{H315} from which it is derived. This indicates that V-J recombination is important in generating diversity within the third hypervariable region of heavy chains. We have identified another J_H segment (J_{HA4}), located 336 nucleotides 3′ to the rearranged J_{H76} segment. This J_H segment is expressed in the heavy chains of anti-levan myeloma proteins, which have truncated third hypervariable regions. We propose that the nucleotide sequence 5′ to J_{HA4} is important for generating V region genes with shortened third hypervariable regions.**

The immunoglobulin gene family is the only eukaryotic gene system found so far in which somatic rearrangements take place during cell differentiation. For the Ig light (L) chain system, it is well established that during lymphocyte development a particular variable (V) and constant (C) gene, which are distant in the germline, are brought into proximity to produce an active gene (1, 2). More detailed analysis has revealed that the V_L gene itself consists of two separate DNA segments, one specifying the NH_2-terminal 95 or 97 amino acids of a classical V_κ or V_λ region, and another, now termed "joining" (J), region specifying the remaining 13 amino acids (3–6). In the germline, the J segment is adjacent to the C gene but separated from it by an intervening sequence. During the somatic rearrangement of V and C genes, a recombination event joins the V and J genes, without altering the spacing between J and C (1, 3). The intervening sequence between VJ and C appears to be transcribed and removed from the pre-mRNA by RNA excision and splicing (7). The J region gene and its flanking DNA sequences thus seem to be involved in two key processes essential for the expression of an immunoglobulin gene: site-specific V-J recombination and VJ-C RNA splicing.

Gene rearrangements in the heavy (H) chain system are of particular biological interest because a single V_H region can be associated with different C_H classes in the same cell or cell lineage (8, 9). Recent work indicates that rearrangement of V_H and J_H is in some respects similar to V_L-J_L joining. During rearrangement, a V_H gene can recombine with one of at least two J_H genes, which lie 5′ to the Cμ gene (10). Early *et al.* (10) have, however, proposed that a third genetic element, D, which encodes amino acids within the third hypervariable region, is recombined between V and J, and have proposed a mechanism for V/D/J recombination.

To study the structure and arrangement of J_H genes and their flanking DNA sequences, we have recently cloned a fragment of chromosomal DNA containing the active V_H + J_H gene from the IgM-secreting mouse plasmacytoma HPC76 (unpublished). This cloned fragment is contiguous in the genome with our Cμ clone, previously isolated from HPC76 (11). In this paper we report the nucleotide sequence of the V_H region and the 5′ half of the C region of the HPC76 μ mRNA. We also report the sequence for the active V_{H76} + J_{H76} gene and the 5′ end of the active Cμ gene. Comparison of the gene sequences with the sequence of the mRNA allowed us to determine the exact boundaries of the intervening sequence between the V_{H76} + J_{H76} and the Cμ genes. We have also determined the sequence of another J_H gene (J_{HA4}), which lies 3′ to the active V_{H76} + J_{H76} gene. Comparison of the J sequences reported here with the germline J_H sequences recently described (10) indicates that V-J recombination is important in generating V_H diversity. Moreover, we find that the sequence 5′ to J_{HA4} has two presumptive recognition sites for V-J recombination. We propose that this allows J_{HA4} to recombine directly with V, omitting a D region, and hence creating a V_H gene with a truncated third hypervariable region.

MATERIALS AND METHODS

Labeling of Restriction Endonuclease Fragments. 5′-Terminal labeling using polynucleotide kinase was as described (4) 3′-Terminal labeling using *Escherichia coli* DNA polymerase I was as follows. Restriction fragments with 5′ overhanging ends (*Hin*dIII, *Hin*fI, *Hpa* II, *Bam*HI, and *Bgl* II) were labeled by "filling in" the 3′ termini. Reaction mixtures (generally 25 μl) contained 10–20 μg of digested DNA in 50 mM Tris·HCl at pH 8, 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 2 μM [α-^{32}P]dCTP for *Hpa* II fragments, or 2 μM [α-^{32}P]dATP for the other fragments, and the other unlabeled dNTPs (each 2 μM). The labeling reaction (37°C, 15 min) was followed by a "chase" (20°C, 15 min) with 50 μM dATP, dCTP, dGTP, and dTTP. Flush-ended *Sma* I fragments were labeled in a similar reaction at the 3′-terminal deoxycytidylate residue by taking advantage of the 3′-exonuclease activity of DNA polymerase I (12). Fragments labeled at one end were then produced by cleavage with another restriction endonuclease or by strand separation on polyacrylamide gels.

DNA Nucleotide Sequence Analysis. The partial chemical degradation method was essentially as described (13) except for the A+G reaction: 5 μl of labeled DNA and 1 μg of calf thymus DNA were incubated with 25 μl of formic acid (20°C, 5 min) and the reaction was terminated by addition of 200 μl

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: V, variable; C, constant; J, joining region; H, heavy; L, light; bp, base pair(s); kb, kilobase pair(s); the symbols C, T, A, and G are used for deoxyribonucleotides in DNA sequences.

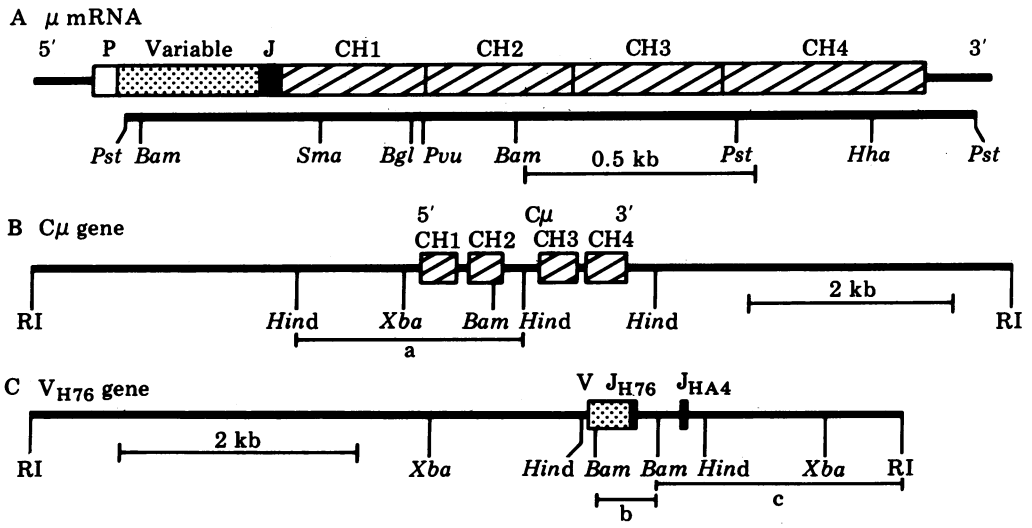

FIG. 1. Restriction endonuclease maps of: (*A*) the μ cDNA insert in the clone pH76μ17 (14); (*B*) the region around the rearranged Cμ gene in clone Ch-H76μ1 (11); and (*C*) the region around the rearranged V_{H76} gene in clone Ch-H76μ119. The left- and rightmost *Pst* I sites of the cDNA clone were generated by the oligo(dG-dC) tailing procedure used to insert the cDNA into the plasmid vector (14, 15). Not all of the *Xba* I or *Hin*dIII sites in Ch-H76μ1, or all of the *Hin*dIII sites in Ch-H76μ119 are shown. In the plasmacytoma HPC76, the *Eco*RI fragments in clones Ch-H76μ1 and Ch-H76μ119 abut, such that the rightmost *Eco*RI site of Ch-H76μ119 is contiguous with the leftmost *Eco*RI site of Ch-H76μ1. The Ch-H76μ1 *Hin*dIII fragment a and the Ch-H76μ119 *Bam*HI and *Bam*HI–*Eco*RI fragments b and c were subcloned in pBR322. kb, Kilobase.

of 1 M sodium acetate containing tRNA at 5 μg/ml. All samples were precipitated with ethanol three times and washed once with 95% (vol/vol) ethanol prior to incubation with 1 M piperidine (90°C, 30 min).

Biological and Physical Containment. Propagation of recombinant DNA molecules was conducted in facilities classified CII and CIII by the Australian Academy of Science Committee on Recombinant DNA (ASCORD) (both classified P3 on National Institutes of Health guidelines), using EK2 host-vector systems, in compliance with both ASCORD and National Institutes of Health guidelines.

RESULTS

Nucleotide Sequence of a μ Chain mRNA. We have previously constructed a recombinant plasmid (pH76μ17, see Fig. 1*A*) that bears a nearly full-length cDNA copy of the μ chain mRNA isolated from the mouse plasmacytoma HPC76 (11, 14). Fig. 2*A* presents a detailed restriction endonuclease map of the

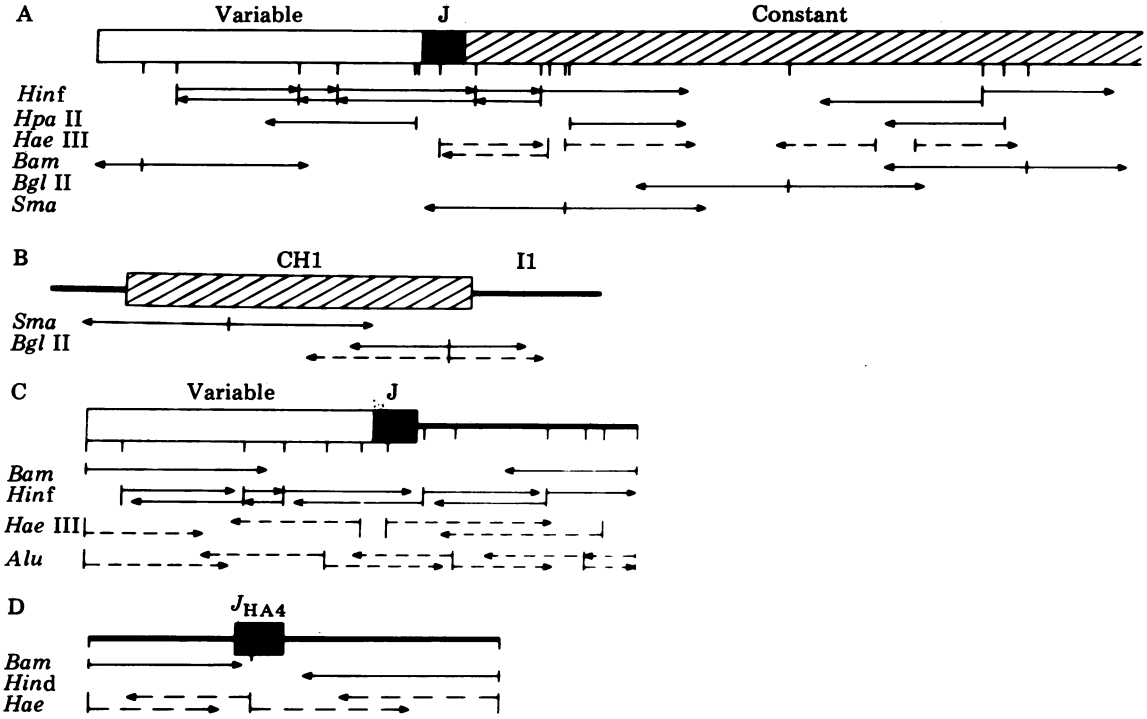

FIG. 2. Strategies for nucleotide sequence analysis. The arrows indicate the direction and extent of sequence determinations. Solid arrows indicate fragments labeled at 3′ termini and broken arrows, fragments labeled at 5′ termini. (*A*) The 5′ half of the pH76μ17 cDNA insert. (*B*) The CH1 domain and flanking intervening sequences of the Cμ gene in Ch-H76μ1. The sequence was determined on the subcloned *Hin*dIII fragment (a in Fig. 1). (*C*) The V + J gene in Ch-H76μ119. The sequence of the subcloned *Bam*HI fragment (b in Fig. 1) was determined. (*D*) The J_{HA4} segment in Ch-H76μ119. The sequence of the *Bam*HI–*Hin*dIII fragment of the *Bam*HI–*Eco*RI subclone (c in Fig. 1) was determined.

cDNA insert and indicates the strategy used for nucleotide sequence analysis.

The sequence determined for the 5′ half (with respect to the mRNA) of the cDNA molecule is presented in Fig. 3, together with the amino acid sequence predicted for the μ polypeptide. The predicted amino acid sequence extends to residue 313 of the C region (numbering sequentially), which is towards the end of CH2. With the exception of three residues (marked by asterisks in Fig. 3), the predicted C region sequence agrees exactly with that of the MOPC 104E μ chain (17). We do not believe that the differences result from nucleotide sequencing error or from an aberration of the cDNA cloning process. In the case of Ser-120 and His-217, the same nucleotides were found within the genomic Cμ sequence that is apparently expressed within this plasmacytoma (see below), and for Glu-225, two base changes would be required to convert this codon to that for the His residue proposed by Kehry *et al.* (17), and the nucleotide sequencing appears unambiguous. Because the Cμ gene occurs as a single copy (18), these differences might reflect a hitherto undetected polymorphism of the BALB/c Cμ gene, or may reflect amino acid sequencing error.

The V region is homologous with other known V_H amino acid sequences. Within the first three framework regions there are only 16 (out of 69) amino acid residues that differ from the McPC 603 sequence (10). However, in addition to the 22 base changes giving rise to these amino acid differences, there are 20 silent base differences from the McPC 603 nucleotide sequence (10).

Sequence of the First Coding Region of a Rearranged Cμ Gene. From DNA of the plasmacytoma HPC76, we have previously cloned a 9.9-kb *Eco*RI fragment bearing a Cμ gene that has undergone somatic rearrangement (11). This is the Cμ gene that is active in this plasmacytoma (S. Cory, personal communication). Fig. 1*B* shows a restriction endonuclease cleavage map of this clone (Ch-H76μ1) and shows that the Cμ gene is split into four coding segments (CH1–CH4), each of which encodes a polypeptide domain (11).

To define the exact 5′ terminus of the Cμ gene, we determined the nucleotide sequence of the first coding region (CH1), using the strategy illustrated in Fig. 2*B*. The boundaries of the coding region were deduced by comparing the nucleotide sequences of the cDNA and the Cμ gene (see Fig. 3). CH1 begins (at the 5′ end) with the first nucleotide of the codon for Ser-115. The coding region is identical with the cDNA sequence up to the first nucleotide of the codon for Ala-222, which is followed by the first intervening sequence (denoted I1, see Fig. 3). There are no intervening sequences within CH1, a possibility that had not been eliminated by previous restriction mapping studies (11).

Sequence of the Active V_H Gene. In work to be described elsewhere, we have cloned the *Eco*RI fragment of chromosomal DNA from HPC76 that bears the V_H gene that is active in this plasmacytoma. Fig. 1*C* shows a restriction map of the clone Ch-H76μ119, which bears a V_{H76} as well as a J_{H76} sequence. In the genome, the *Eco*RI fragment of Ch-H76μ119 lies immediately to the left of that in the Cμ clone (S. Cory, personal communication).

Hybridization experiments showed that the V_{H76} and J_{H76} sequences were within the 0.5-kb *Bam*HI fragment (b in Fig. 1*C*), and therefore, to determine their structure, we sequenced that fragment, using the strategy shown in Fig. 2*C*. The nucleotide sequence of the genomic V gene and its predicted amino acid sequence are presented in Fig. 4. The sequence starts at the left-hand *Bam*HI site (Fig. 1*C*), which corresponds to amino acid residues 16 and 17 of the V region. The subsequent sequence is identical to that of the V region in the mRNA

```
                                   10                                      20                                      30
                          GluSerGlyGlyGlyLeuValGlnProGlyGlySerMetLysLeuSerCysValAlaSerGlyPheThrPheSerAsnTyrTrp
cDNA              GGGGGGGGAGGAGTCTGGAGGAGGCTTGGTGCAACCTGGAGGATCCATGAAACTCTCCTGTGTTGCCTCTGGATTCACTTTCAGTAACTACTGG

              40                                      50                                      60                                      70
     MetAsnTrpValArgGlnSerProGluLysGlyLeuGluTrpValAlaGluIleArgLeuLysSerGlyTyrAlaThrHisTyrAlaGluSerValLysGlyArgPheThrIleSerArg
cDNA ATGAACTGGGTCCGCCAGTCTCCAGAGAAGGGGCTTGAGTGGGTTGCTGAAATTAGATTGAAATCTGGTTATGCAACACATTATGCGGAGTCTGTGAAAGGGAGGTTCACCATCTCAAGA

              80                                      90                                      100                                     110
     AspAspSerLysSerSerValTyrLeuGlnMetAsnAsnLeuArgAlaGluAspThrGlyIleTyrTyrCysThrArgProGlyValProAspTyrTrpGlyGlnGlyThrThrLeuThr
cDNA GATGATTCCAAAAGTAGTGTCTACCTGCAAATGAACAACTTAAGAGCTGAAGACACTGGCATTTATTACTGTACCAGGCCGGGGGTCCCCGACTACTGGGGCCAAGGCACCACTCTCACA
GENE                                                                                                   AGATTCTTACCCCTCCATTTCTCTTTTATCCCTCTC
                                                                                                                          IVS
        V<--|-->C     *
              120                                     130                                     140                                     150
     ValSerSerGluSerGlnSerPheProAsnValPheProLeuValSerCysGluSerProLeuSerAspLysAsnLeuValAlaMetGlyCysLeuAlaArgAspPheLeuProSerThr
cDNA GTCTCCTCAGAGAGTCAGTCCTTCCCAAATGTCTTCCCCCTCGTCTCCTGCGAGAGCCCCCTGTCTGATAAGAATCTGGTGGCCATGGGCTGCCTGGCCCGGGACTTCCTGCCCAGCACC
GENE TGGTCCTCAGAGAGTCAGTCCTTCCCAAATGTCTTCCCCCTCGTCTCCTGCGAGAGCCCCCTGTCTGATAAGAATCTGGTGGCCATGGGCTGCCTGGCCCGGGACTTCCTGCCCAGCACC

              160                                     170                                     180                                     190
     IleSerPheThrTrpAsnTyrGlnAsnAsnThrGluValIleGlnGlyIleArgThrPheProThrLeuArgThrGlyGlyLysTyrLeuAlaThrSerGlnValLeuLeuSerProLys
cDNA ATCTCCTTCACCTGGAACTACCAGAACAACACTGAAGTCATCCAGGGTATCAGAACCTTCCCAACACTGAGAACAGGGGGCAAGTACCTAGCCACCTCACAAGTGTTGCTGTCTCCCAAG
GENE ATCTCCTTCACCTGGAACTACCAGAACAACACTGAAGTCATCCAGGGTATCAGAACCTTCCCAACACTGAGAACAGGGGGCAACTACCTAGCCACCTCACAAGTGTTGCTGTCTCCCAAG

              200                                     210                         *           220                *                    230
     SerIleLeuGluGlySerAspGluTyrLeuValCysLysIleHisTyrGlyGlyLysAsnArgAspLeuHisValProIleProAlaValAlaGluMetAsnProAsnValAsnValPhe
cDNA AGCATCCTTGAAGGTTCAGATGAATACCTGGTATGCAAAATCCACTACGGAGGCAAAAACAGAGATCTGCATGTGCCCATTCCAGCTGTCGCAGAGATGAACCCCAATGTAAATGTGTTC
GENE AGCATCCTTGAAGGTTCAGATGAATACCTGGTATGCAAAATCCACTACGGAGGCAAAAACAGAGATCTGCATGTGCCCATTCCAGGTAAGAACCAAACCCTCCCAGCAGGTGCCCAGGC
                                                                                                                         I1
              240                                     250                                     260                                     270
     ValProProArgAspGlyPheSerGlyProAlaProArgLysSerLysLeuIleCysGluAlaThrAsnPheThrProLysProIleThrValSerTrpLeuLysAspGlyLysLeuVal
cDNA GTCCCACCACGGGATGGCTTCTCTGGCCCTGCACCACGCAAGTCTAAACTCATCTGCGAGGCCACGAACTTCACTCCAAAACCGATCACAGTATCTTGGCTAAAGGATGGGAAGCTCGTG

              280                                     290                                     300                                     310
     GluSerGlyPheThrThrAspProValThrIleGluAsnLysGlySerThrProGlnThrTyrLysValIleSerThrLeuThrIleSerGluIleAspTrpLeuAsnLeuAsnValTyr
cDNA GAATCTGGCTTCACCACAGATCCGGTGACCATCGAGAACAAAGGATCCACACCCCAAACCTACAAGGTCATAAGCACACTTACCATCTCTGAAATCGACTGGCTGAACCTGAATGTGTAC
```

FIG. 3. Comparison of the nucleotide sequence of the μ cDNA and the Cμ gene. The sequence of the mRNA sense strand of the cDNA clone pH76μ17 is shown, together with the amino acid sequence predicted for the μ polypeptide. The amino acid sequence has been aligned to place the first Cys residue at position 22 (16). The G residues at the start of the sequence result from the oligo(dG·dC) tailing procedure used to insert the cDNA into the plasmid vector (14, 15). The sequence of the mRNA sense strand of the first coding region (CH1) of the Cμ gene in the clone Ch-H76μ1 is shown below the cDNA sequence. Intervening sequences (IVS and I1) are underlined. Residues in the C region that differ from the previously published μ chain amino acid sequence (17) are marked (*). The V/C junction is defined in the text.

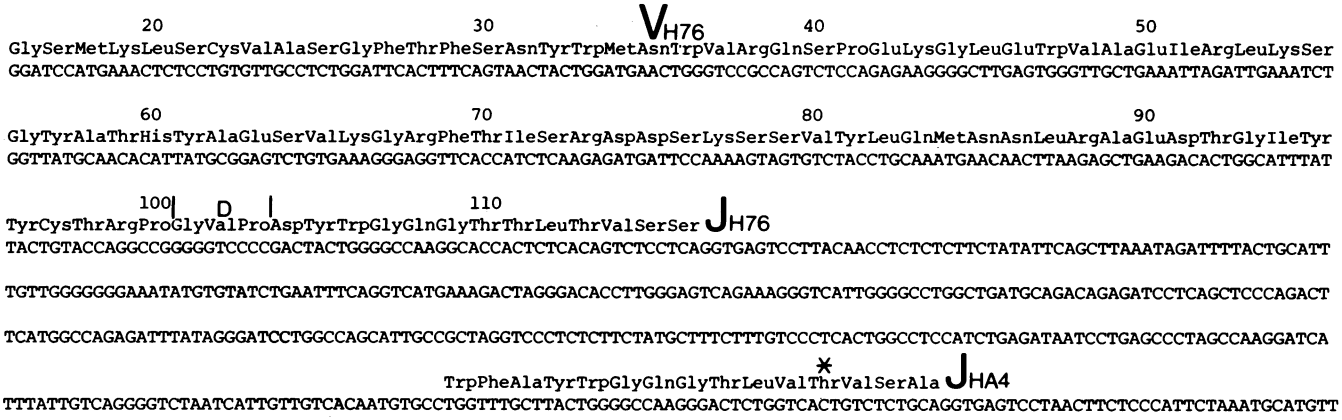

FIG. 4. Nucleotide sequence of the active V_{H76} + J_{H76} gene and a neighboring J_H segment. The sequence of the mRNA sense strand from the clone Ch-H76μ119 is shown, together with the amino acid sequence predicted for the rearranged V gene (with the same numbering system as in Fig. 3). The boundaries of J_{H76} are placed by comparison with the unrearranged J_H segment (J_{H315}, ref. 10) from which this sequence is derived (see text). The D region is defined by Schilling *et al.* (19) and Early *et al.* (10). The region 3′ to the V + J gene was examined in all translational reading frames. J_{HA4} was the only segment found that could encode any of the known J_H amino acid sequences (19). The asterisk marks a residue in J_{HA4} that has apparently undergone somatic mutation (to Pro) in one of the anti-levan V_H regions (see *Results*).

to the codon for Ser-116, without interruption, and diverges one nucleotide thereafter. Because the Cμ gene could encode from Ser-115, the rearranged V and C genes contain all of the information required to encode the μ mRNA. Thus, as in the κ and λ systems (3, 4), the translocated V_H gene is contiguous with its J_H segment, in agreement with the recent results of Early *et al.* (10).

There are seven nucleotides at the 3′ end of the V gene (indicated in italics) that are repeated at the 5′ end of the C gene:

mRNA ACAGTC*TCCTCAG*AGAGT
V gene ACAGTC*TCCTCAG*GTGAG
C gene CTCTGG*TCCTCAG*AGAGT.

Terminally redundant sequences appear to be a common feature of coding regions split by intervening sequences (4). In most cases, intervening sequences are removed by a RNA excision and splicing event that occurs 5′ to a GT and 3′ to an AG sequence (20). We therefore speculate that splicing occurs in the pre-mRNA at the ends of the underlined sequences, which are at the 3′ end of the redundant sequence.

What is the sequence of the J_H segment to which the V_{H76} gene is joined? The position of the rearranged V_{H76} gene relative to the Cμ gene places it at the same position as the germline J_{H315} segment whose sequence was recently determined by Early *et al.* (10). Moreover, the last 39 nucleotides of the complete V_{H76} gene (encoding Asp-104 to Ser-116) are identical with the last 39 nucleotides of the J_{H315} segment (see Fig. 5). We therefore conclude that the embryonic V_{H76} gene has recombined with the J_{H315} segment. There are, however, two codons (for Tyr-Phe) at the beginning of J_{H315} (Fig. 5) that are not within the rearranged V_{H76} sequence (see *Discussion*).

Early *et al.* (10) and Schilling *et al.* (19) have suggested that a part of the third hypervariable region is encoded by neither the embryonic V_H gene nor the J_H gene and have proposed that a third gene segment, D, is required for the creation of an active V_H gene. In the HPC76 heavy chain, the amino acid residues that correspond to the D region are Gly-101, Val-102, and Pro-103 (Fig. 3). The nucleotide sequence encoding these residues, GGGGTCCCC, is a perfect palindrome, as has been noted for another D region (10). This sequence is not part of the embryonic J_{H76} segment, nor is it present in the preceding 577 nucleotides (10).

A Nonrearranged J_H Segment Downstream from the Active V_{H76} Gene. We determined the nucleotide sequence downstream from the V_{H76} gene (fragment c in Fig. 1*C*) by using the strategy indicated in Fig. 2*D*, and we found a second J_H segment located 336 nucleotides 3′ to the end of the active V_{H76} gene (Fig. 4). The amino acid sequence encoded by this J is found in V_H regions of three anti-levan myeloma proteins, A4, A47N, and E109 (21), and thus we designated it J_{HA4}. None of these V_H regions, however, contains the first amino acid (Trp) potentially encoded by J_{HA4} (see *Discussion*). In another anti-levan V_H region, U61, the Thr residue indicated by an asterisk is replaced by Pro (21). This presumably reflects somatic mutation of the germline J_{HA4} region; only one base change

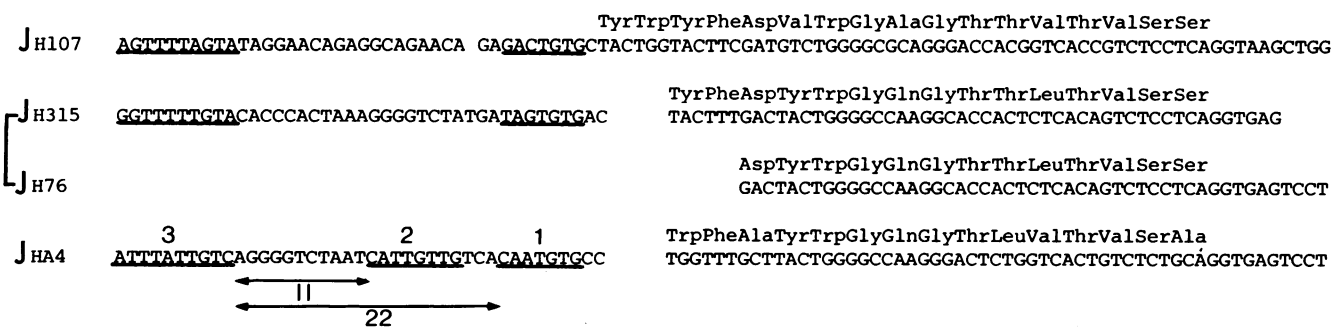

FIG. 5. Comparison of different J_H sequences. The J_{H107} and J_{H315} sequences are taken from Early *et al.* (10). J_{H76} is derived from the embryonic J_{H315}, by recombination with the V_H gene expressed in the plasmacytoma HPC76. Both the coding regions and the 5′ flanking sequences have been aligned to maximize homology. Because the J segments are of different lengths, this results in gaps at the beginning of the coding regions. Presumptive recombination recognition sites are underlined (5, 6, 10). The sites (1, 2, and 3) ahead of J_{HA4} are discussed in the text. The distances 11 and 22 base pairs (bp) are discussed in the text.

the fact that binding activity is lost after the plasma membrane is solubilized (8). EGF receptors, however, were detected as a chemically crosslinked complex with ^{125}I-labeled EGF and their subunit molecular weights were estimated to be in the range 180,000–195,000 (6, 8, 24). By definition, the specific cell surface components to which radioactive polypeptide hormones bind are hormone receptors (25). Therefore, the lack of EGF binding in mouse A9 cells must be correlated with a deficiency in receptors or components involved in synthesizing functional receptors. The simplest and most favorable interpretation of our results is that human chromosome 7 carries a gene(s) for the receptor structural protein.

It has been postulated that membrane proteins embedded in the lipid bilayer are synthesized on membrane-bound ribosomes (26). During polypeptide synthesis the ribosomes are directed to specific sites in the endoplasmic reticulum and the growing polypeptide chain is extruded through the membrane. Carbohydrate is then added to the proteins, and the completed glycoproteins finally appear in the plasma membrane. It is therefore possible that the A9 cells are capable of producing EGF receptors but they are either not exposed at the cell surface or they are exposed but the carbohydrate moiety, presumably necessary for EGF binding, is lacking. We have tested possible accumulation of premature receptor proteins with negative results—no binding of ^{125}I-labeled EGF was found in A9 cells fixed with ethanol and treated with acetone or dilute Triton X-100 (unpublished results).

Not only a defect in glycosylation of the receptor proteins but also a defect in the biosynthesis of amino sugars could lead to incomplete glycosylation of receptor proteins and to decreased exposure to the cell surface (27). Such defects would probably result in decreased formation of complex glycolipids and this may indirectly contribute to diminished EGF-binding ability. Alternatively, the full function of the EGF receptors may require the presence of normal surface components in the vicinity of the receptor that facilitate EGF binding in a stereospecific manner. If this is the case for the deficiency in A9, these components should be encoded by the gene(s) on human chromosome 7.

Another possibility relates to the findings that sarcoma virus-transformed cells release a family of EGF-like growth factors into culture fluid and that they compete for EGF receptors (21, 28). A9 cells are mutants derived from mouse L cells which are known to secrete nerve growth factor-like polypeptides (29). A9 cells are not known for the presence of murine sarcoma virus although they possess many features that are commonly associated with transformed cells (11). Thus, the apparent loss of EGF-binding ability could be attributed to the blocking or masking of EGF receptors by closely related polypeptides. Then, restoring of binding ability in hybrid cells could be considered to be the result of suppression of polypeptide synthesis by human genes on chromosome 7. This interesting possibility must be extensively examined although our preliminary experiments indicated that culture medium of A9 cells does not inhibit the binding of ^{125}I-labeled EGF to human HeLa cells (unpublished results).

In summary, we have demonstrated the expression and segregation of a polypeptide hormone-binding ability on the cell surface of human–mouse cell hybrids. This is an important prerequisite for identifying a number of genes that control a complex hormonal action *in vitro*. It will be of great importance to demonstrate the biological response of hybrid cells after EGF treatment. We have observed that the receptors expressed in the C3B4 hybrid line are subject to "down-regulation" (30) and the cell-bound EGF is internalized (unpublished results). By using a number of mutants with different lesions one would be able to dissect the cascade of EGF/receptor-mediated hormonal signal transfer mechanisms. Considering the complexity as discussed above, we use a gene symbol *EGFS* to represent the human gene(s) that complements the lesion of mouse A9 cells in EGF binding and that is involved in determining *sensitivity* to the EGF interaction. *EGFS* is located on the p22–qter region of human chromosome 7.

We thank Drs. H. V. Aposhian and F. H. Ruddle for their continuous interest and encouragement throughout this work. This work was supported by National Institutes of Health Grant GM 24375. N.S. is a recipient of American Cancer Society Junior Faculty Research Award JFRA-9.

1. Carpenter, G. & Cohen, S. (1978) in *Biochemical Actions of Hormones*, ed. Litwack, G. (Academic, New York), Vol. 5, pp. 203–247.
2. Gospodarowicz, D. & Moran, J. S. (1976) *Annu. Rev. Biochem.* **45**, 531–558.
3. Carpenter, G. & Cohen, S. (1979) *Annu. Rev. Biochem.* **48**, 193–216.
4. Gospodarowicz, D., Greenburg, G., Bialecki, H. & Zetter, B. R. (1978) *In Vitro* **14**, 85–118.
5. Schlessinger, J., Shecter, Y., Willingham, M. C. & Pastan, I. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2659–2663.
6. Das, M. & Fox, F. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2644–2648.
7. Carpenter, G. & Cohen, S. (1977) *Biochem. Biophys. Res. Commun.* **79**, 545–552.
8. Hock, R. A., Nexø, E. & Hollenberg, M. D. (1979) *Nature (London)* **277**, 403–405.
9. Shimizu, N., Behzadian, M. A. & Shimizu, Y. (1979) *Cytogenet. Cell Genet.*, **25**, 201–202.
10. Hayflick, L. (1965) *Exp. Cell Res.* **37**, 614–636.
11. Littlefield, J. W. (1966) *Exp. Cell Res.* **41**, 190–196.
12. Littlefield, J. W. (1964) *Science* **145**, 709–710.
13. Siminovitch, L. (1976) *Cell* **7**, 1–11.
14. Sauer, F., Greenstein, R. M., Reardon, P. & Riddick, D. (1977) *Obstet. Gynecol.* **49**, 101–104.
15. Nichols, E. A. & Ruddle, F. H. (1973) *J. Histochem. Cytochem.* **21**, 1066–1081.
16. Shimizu, N., Giles, R. E., Kucherlapati, R. S., Shimizu, Y. & Ruddle, F. H. (1977) *Somat. Cell Genet.* **3**, 47–60.
17. Shimizu, N., Shimizu, Y. & Ruddle, F. H. (1978) *Cytogenet. Cell Genet.* **22**, 441–445.
18. Shimizu, N., Shimizu, Y., Kucherlapati, R. S. & Ruddle, F. H. (1976) *Cell* **7**, 123–130.
19. Kozak, C. A., Lawrence, J. B. & Ruddle, F. H. (1977) *Exp. Cell Res.* **105**, 109–117.
20. Carpenter, G. & Cohen, S. (1976) *J. Cell Biol.* **71**, 159–171.
21. Todaro, G. J., De Larco, J. E. & Cohen, S. (1976) *Nature (London)* **264**, 26–31.
22. McKusick, V. A. & Ruddle, F. H. (1977) *Science* **196**, 390–405.
23. Kahan, B. & DeMars, R. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1510–1514.
24. Das, M., Miyakawa, T., Fox, C. F., Pruss, R. M., Aharonov, A. & Herschman, H. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2790–2794.
25. Roth, J. (1975) *Methods Enzymol.* **37B**, 66.
26. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67**, 852–862.
27. Pratt, R. M. & Pastan, I. (1978) *Nature (London)* **272**, 68–70.
28. De Larco, J. E. & Todaro, G. J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4001–4005.
29. Oger, J., Arnason, B. G. W., Pantazis, N., Lehrich, J. & Young, M. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1554–1558.
30. Aharonov, A., Pruss, R. M. & Herschman, H. R. (1978) *J. Biol. Chem.* **253**, 3970–3977.