Proc. Natl. Acad. Sci. USA
Vol. 78, No. 1, pp. 509–513, January 1981
Immunology

Sequences of the joining region genes for immunoglobulin heavy chains and their role in generation of antibody diversity

(rearranged DNA/γ_{2a} mRNA/codon deletion/somatic mutation/allelic exclusion)

NICHOLAS M. GOUGH* AND ORA BERNARD†

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, September 17, 1980

ABSTRACT **To assess the contribution to immunoglobulin heavy chain diversity made by recombination between variable region (V_H) genes and joining region (J_H) genes, we have determined the sequence of about 2000 nucleotides spanning the rearranged J_H gene cluster associated with the V_H gene expressed in plasmacytoma HPC76. The active V_{H76} gene has recombined with the second germ-line J_H gene. The region we have studied contains two other J_H genes, designated J_{H3} and J_{H4}. No other J_H gene was found within the region 1000 nucleotides downstream from J_{H4}. Between J_{H3} and J_{H4} there is a pseudo-J_H sequence with substantial homology to the authentic J_H genes. The four J_H genes whose sequences now are known can account for all known J_H amino acid sequences. The J_H genes are more divergent than the J_κ genes and vary in length, encoding either 15 or 17 amino acid residues. Because J_H regions comprise part of the third hypervariable region (HV3), combinatorial V_H–J_H joining substantially augments V_H diversity. Moreover, a V_H gene can recombine with each J_H gene at several positions, and either one or two germ-line J_H codons can be excised. This J_H truncation markedly reduces the length of *HV3* and hence must alter antigen-binding specificity. We have also determined the sequence of the J_{H4} region in two different γ_{2a} mRNAs and have found that each has suffered a point mutation (aspartate to asparagine) which would alter the charge of the antigen-binding site.**

The genes that encode the variable (V) and the constant (C) portions of immunoglobulin (Ig) molecules are distant in germ-line DNA and, during lymphocyte development, are brought into proximity to produce an active gene (1–5). Light chain *V* genes actually consist of two separate DNA segments, one specifying the NH_2-terminal 95 or 97 amino acids of a classical V_λ or V_κ region and the other [the "joining" (J) region] specifying the remaining 13 amino acids (6–9). The J_L genes are near to their respective C_L genes but separated from them by intervening sequences. During lymphocyte development, a *V* gene is joined to a *J* gene without altering the spacing between *J* and *C*, the DNA originally separating the *V* and *J* genes being deleted (9). The intervening sequence between *V–J* and *C* appears to be transcribed and then excised from the pre-mRNA (10, 11).

In addition to playing a necessary role in *V* gene rearrangement and mRNA processing, the four J_κ genes are involved to some extent in the generation of antibody diversity. The first J_κ amino acid residue is on the border of the third hypervariable region (*HV3*)‡ and therefore amino acid substitutions at this position may alter the properties of the antigen-binding site. Thus, combinatorial association of different V_κ genes with the four J_κ genes, which differ at the first codon (8, 9), generates different sequences in or near the complementarity-determining region. Moreover, it has been proposed, on the basis of V_κ amino acid sequences, that the first J_κ residue can be altered by *V–J* recombination, allowing some further variation (8, 9, 12).

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.

Recently, three heavy chain *J* genes, located ≈8.5–9.0 kilobases (kb) 5′ to the C_μ gene, have been identified (13, 14). Early *et al.* (13) analyzed two J_H genes in germ-line (sperm) DNA, whereas we have studied the J_H cluster downstream from the active V_H gene in an IgM-producing plasmacytoma (14). In this rearranged J_H cluster the expressed V_H gene has recombined with the second of the two J_H genes studied by Early *et al.*, thereby deleting the first.

To establish the full repertoire of J_H genes, we have extended our analysis of the rearranged J_H locus to a total of 2.2 kb downstream from the active V_H gene. We have identified a fourth J_H gene within this sequence, as well as a sequence that appears to be a pseudo-J_H gene. We have also determined the nucleotide sequence of the *J* loci in two γ_{2a} mRNAs. We show that somatic diversification occurs by choice of the J_H gene to be activated, truncation of the third hypervariable region during *V–J* recombination, and somatic base substitution.

MATERIALS AND METHODS

We have previously cloned the $V_H + C_\mu$ gene that is expressed in the IgM-producing plasmacytoma HPC76 (5, 15) (Fig. 1*a*). The indicated DNA fragments from this clone were labeled at their 5′ or at their 3′ end as described (14). The partial chemical degradation of Maxam and Gilbert was used (16).

RESULTS

Nucleotide Sequence of the J_H Locus. We determined the nucleotide sequence of the J_H locus downstream from the active V_H gene of plasmacytoma HPC76 by using the strategy in Fig. 1*c* and in ref. 14. The sequence starts near the end of the V region and extends 2.2 kb downstream (Fig. 2) The J_{H76} sequence is derived from the germ-line J_H gene here designated J_{H2} (13, 14). Downstream from J_{H2} we previously identified (14) a J_H gene (J_{H3} in Fig. 2) that encodes the J region in three myeloma proteins with anti-levan activity (17). To locate other J_H genes we searched the nucleotide sequence downstream from V_{H76} in all translational reading frames to identify sequences encoding the known J_H amino acid sequences (18). Because there might be J_H genes encoding amino acid sequences not yet identified, we also searched for the short conserved sequences 5′ to the known J_H genes (13, 14). One further J_H gene was identified: J_{H4}, located ≈520 nucleotides downstream from J_{H3}. We also

Abbreviations: V, variable; C, constant; J, joining region; kb, kilobase(s); bp, base pair(s).

* Present address: European Molecular Biology Laboratory, Postfach 10.2209, 6900 Heidelberg, Federal Republic of Germany.
† To whom reprint requests should be addressed.

‡ Kabat, E. A., Wu, T. T. & Bilofsky, H. (1979) *Sequence of Immunoglobulin Chains*, National Institutes of Health publication number 80-2008.

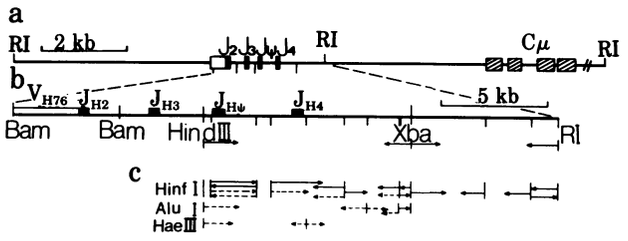

FIG. 1. Strategies for nucleotide sequence analysis of the J_H gene cluster in HPC76. (*a*) Restriction endonuclease map of cloned *Eco*RI (RI) fragments bearing the active V_H–J_H–C_μ locus from plasmacytoma HPC76 (5, 15). (*b*) The *Bam*HI and *Bam*HI/*Eco*RI fragments were subcloned in pBR322. (*c*) Arrows indicate the direction and extent of sequence determinations. Solid arrows indicate fragments labeled at 3′ termini (14) and broken arrows indicate those labeled at 5′ termini (6). The *Bam*HI/*Eco*RI subclone was digested with *Hin*dIII, *Xba* I, and *Eco*RI; the *Hin*dIII/*Xba* I and *Xba* I/*Eco*RI fragments were separated on a 5% acrylamide gel and digested with the restriction endonucleases indicated. There are three short gaps where we were unable to obtain fragments sufficiently pure for accurate sequence determination.

identified a sequence ($J_{H\psi}$ in Fig. 2) that bears strong homology with other J_H genes and which we propose represents a pseudo-J_H gene (see *Discussion*).

Nucleotide Sequence Within Two γ_{2a} mRNAs. The J_{H4} gene encodes the J region found in the published MOPC173 γ_{2a} amino acid sequence (19). To confirm the published amino acid sequence, we determined the nucleotide sequence of the *J* region and the 5′ portion of the *C* region of cloned cDNA copies of both the MOPC173 and the HOPC1 γ_{2a} mRNA (Fig. 3). The two γ_{2a} nucleotide sequences are identical throughout the region examined, with the exception of one nucleotide within the codon for amino acid residue 173: in MOPC173 a leucine residue is predicted, and in HOPC1, a valine. The difference, which gives rise to a *Pvu* II restriction site in the HOPC1 sequence, cannot be due to an aberration of cDNA cloning because two independent MOPC173 γ_{2a} cDNA clones lack this *Pvu* II site and four HOPC1 γ_{2a} cDNA clones contain it (20). Because there probably is only one $C_{\gamma 2a}$ gene per haploid mouse genome (21), the difference between the two mRNAs may reflect a residual polymorphism of the γ_{2a} gene in BALB/c mice or a mutation within one of the tumor lines.

As indicated in Fig. 3, our MOPC173 sequence differs from the published amino acid sequence at five positions in the $C_{\gamma 2a}$ region and at one in the J_H region. These differences are unlikely to be due to nucleotide sequencing errors because the HOPC1 mRNA sequence is identical (except for the difference noted), and the nucleotide sequence determination was unambiguous.

A significant finding is that the J region of the MOPC173 and HOPC1 γ_{2a} mRNAs encodes an asparagine residue, whereas the genomic J_{H4} gene in HPC76 encodes an aspartate. Thus, the J regions within two heavy chain mRNAs differ by one nucleotide from the genomic *J* region that encodes them, and therefore there has been a somatic base change.

```
96                VH76                    110
TyrCysThrArgProGlyValProAspTyrTrpGlyGlnGlyThrThrLeuThrValSerSer JH2
TACTGTACCAGGCCGGGGGTCCCCGACTACTGGGGCCAAGGCACCACTCTCACAGTCTCCTCAGGTGAGTCCTTACAACCTCTCTCTTCTATATTCAGCTTAAATAGATTTTACTGCATTTGT

TGGGGGGGAAATATGTGTATCTGAATTTCAGGTCATGAAGGACTAGGGACACCTTGGGAGTCAGAAAGGGTCATTGGGAGCCCTGGCTGATGCAGACAGACATCCTCAGCTCCCAGACTTCAT
GGCCAGAGATTTATAGGGATCCTGGCCAGCATTGCCGCTAGGTCCCTCTCTTCTATGCTTTCTTTGTCCCTCACTGGCCTCCATCTGAGATAATCCTGAGCCCTAGCCAAGGATCATTTATTG
                                  TrpPheAlaTyrTrpGlyGlnGlyThrLeuValThrValSerAla JH3
TCAGGGGTCTAATCATTGTTGTCACAATGTGCCTGGTTTGCTTACTGGGGCCAAGGGACTCTGGTCACTGTCTCTGCAGGTGAGTCCTAACTTCTCCCATTCTAAATGCATGTTGGGGGGAT

TCTGAGCCTTCAGGACCAAGATTCTCTGCAAACGGGAATCAAGATTCAACCCCTTTGTCCCAAAGTTGAGACATGGGTCTGGGTCAGGGACTCTCTGCCTGCTGGTCTGTGGTGACATTAGA

ACTGAAGTATGATGAAGGATCTGCCAGAACTGAAGCTTGAAGTCTGAGGCAGAATCTTGTCCAAGGTCTATCGGACTCTTGTGAGAATTAGGGGCTGACAGTTCATGGTGACAATTTCAGG
                                  JHψ
GTCAGTGACTGTCTGGTTTCTCTGAGGTGAGGCTGGAATATAGGTCACCTTGAAGACTTAAGAGGGGTCCAGGGGGCTTCTGCACAGGCAGGGAACAGAATGTGGAACAATGACTTGAATG
      ~50N                                        TyrTyrAlaMetAspTyrTrpGlyGlnGlyThrSerValThrValSerSer JH4
GTTGG---------GACTCTGGGGTTTTTGTCGGGTATAGAGGAAAAATCCACTATTGTGATTACTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGGTAAGAAT

GGCCTCTCCAGGTCTTTATTTTTAACCTTTGTTATGGAGTTTTCTGAGCATTGCAGACTAATCTTGGATATTTGCCCTGAGGGAGCCGGCTGAGAGAAGTTGGGAAATAAATCTGTCTAGGG

ATCTCAGAGCCTTTAGGACAGATTATCTCCACATCTTTGAAAAACTAAGANTCCCTAGATGATGGGATAGGGACTTTGGAGGCTCATTTGAGGGAGATGCTAAAACAATCCTATGGCTGGAG

GGATAGTTGGGGCTGTAGTTGGAGATTTTCAGTTTTTAGAATGAAGTATTAGCTGCGGCATAGGGACAAAAAGTGGAGTGGGGCACTTTCTTTAGATTTGTGAGGAATGTTCCACACTAGAT

TGTTTAAAACTTCATTTGTTGGAAGGAGCTGTCTTAGTGATTGAGTCAAGGGAGAAAGGCATCTAGCCTCGGTCTCAAAAGGGTAGTTGCTGTCTAGAGAGGTCTGGTGGAGCCTGCAAAAG

TCCAGCTTTCAAAGGAACACAGAAGTATGTGTATGGAATATTAGAAGATGTTGCTTTTACTCTTAAGTTGGTTCCTAGGAAAAATAGTTAAATACTGTGACTTTAAAATGTGAGAGAGGGTT
                                  ~60N
TCAGTACTCATTTTTAAATGTCCAAAATTTTGCA--------CTAAACTTAAAGATTTAACCGAGGAATGGGAGTGAGGCTCTCTCATACCCTATTCAGAACTGACTTTTAACAATAATAAA
                                        ~70N
TTAAGTTTAAAATATTTTTAAATGAATTGAGCAATGTTGAGTTGG-------AGGTAAATAAATTTGTAGCTGTGGTTTGAAGAAGTGGTTTTGAAACACTCTGTCCAGCCCCACCAAACCG

AAAGTCCAGGCTGAGCAAAACACCACCTGGGTAATTTCTAAAATAAGTTGAGGATTCAGCCGAAACTGGAGAGGTCCTCTTTTTAACTTATTGAGTTCAACCTTTTAATTTTAGCTTGAGTA

GTTCTAGTTTCCCCAAACTTAAGTTTATCGACTTCTAAAATGTATTTAGAATTC
```

FIG. 2 Nucleotide sequence of J_H gene cluster in plasmacytoma HPC76. The sequence of the mRNA sense strand is shown together with the amino acid sequence predicted for each J_H gene. The sequence starts from the codon for amino acid 96 of V_{H76} and extends through J_{H2}, J_{H3}, and J_{H4}; it includes about 1000 nucleotides 3′ to the J_H cluster, up to the *Eco*RI site (Fig. 1*b*). The underlined sequences are two common sequences 5′ to each J_H. This sequence has been corrected for typing errors in the region spanning J_{H2} to J_{H3} previously published (14). $J_{H\psi}$ indicates a sequence homologous to the J_H genes and which is postulated to be a pseudo-J_H gene (see text).

FIG. 3. Nucleotide sequence of V_H–J_H–$C_{\gamma 2a}$ junction in MOPC173 and HOPC1 cDNAs. The sequence starts with the codon for amino acid 96 and stops at the codon for amino acid 194. Sequences common to MOPC173 and HOPC1 are indicated by a continuous line. When an amino acid residue predicted by nucleotide sequencing does not agree with the published sequence (19), the published amino acid is indicated above that deduced from the nucleotide sequence. The asterisk indicates the *Bam*HI site, from which the nucleotide sequence was determined.

DISCUSSION

We have determined the nucleotide sequence of 2.2 kb of the rearranged J_H locus in the plasmacytoma HPC76. The sequence contains three J_H genes, the first (13) having been deleted during the recombination event that linked the expressed V_H gene to J_{H2}. No further J_H genes occur 1 kb downstream from J_{H4}. As discussed below, all of the known J_H amino acid sequences can be assigned to these four J_H genes. Moreover, no J_H genes have been found in association with other germ-line C_H genes (unpublished data). We believe therefore that $J_{H1}J_{H4}$ constitute the full set of J_H genes. During preparation of this manuscript we learned of work by Sakano *et al.* (22) which has led to the same conclusion.

The D Segment. It has been proposed (13, 18) that amino acid residues of HV3 NH_2-terminal to the J_H region, the "D_H region," are encoded by an element independent of germ-line V_H and J_H genes. On the basis of amino acid sequences, Schilling *et al.* (18) proposed, for each known V_H region, which residues are encoded by *D* and which by *J*. Because we now know the exact coding capacity of each J_H gene, we have redefined the boundaries of D and J (Table 1). For example, for MOPC173, Schilling *et al.* (18) proposed that the D region is Pro-Tyr and the beginning of J is Tyr-Ala. Because the NH_2-terminal residues encoded by J_{H4} are Tyr-Tyr-Ala, we propose that the MOPC173 D region contains only the proline residue.

***V–J* Recombination and the Generation of Antibody Diversity.** Two types of alterations in hypervariable regions are likely to markedly influence the antigen-binding site (23). First, amino acid substitutions that alter the charge or size of a particular residue can greatly change the binding affinity. Second, insertions and deletions in hypervariable loops play a major role, determining whether the loops form a deep cavity or a shallow groove. Fig. 4 shows that the 5′ ends of J_H regions encode amino acids within HV3 and hence form an integral part of the antigen-binding site.

Combinatorial Joining with Four Germ-Line J_H Segments. In contrast to the J_κ genes, which are all the same length (13 codons) and exhibit little sequence variation, Fig. 4 shows that the four J_H genes differ in length and show more sequence divergence, particularly toward the NH_2 terminus. Whereas for J_κ only the first residue is part of HV3, the first four residues of J_{H2} and J_{H3} and the first six of J_{H1} and J_{H4} are within HV3.‡ Within this portion of the J_H sequences there are no invariant positions. Because each J_H gene contributes a distinct sequence to HV3, each presumably has a unique effect upon the configuration of the antigen-binding site. Thus, combinatorial joining of J_H genes with different V_H (or D_H) genes contributes significantly to the final repertoire of antigen-binding sites.

Somatic Base Substitutions. The sequence of the J_{H4} gene in HPC76 DNA and the J region in two mRNAs (MPOC173 and HOPC1) derived from this J_H gene differ by one nucleotide (see Figs. 2 and 3). Because the J_{H4} sequence in HPC76 DNA is identical with the germ-line J_{H4} sequence (22), it follows that the MOPC173 and HOPC1 sequences have undergone somatic mutation. Somatic mutations in immunoglobulin genes occur during the growth of mouse plasmacytomas (24) and the single base difference within the C region of the two γ_{2a} mRNAs (Fig. 3) may have so arisen. However, mutations probably also occur during the growth and development of normal lymphocytes and those within *V*, *D*, and *J* genes would diversify the antigen-binding repertoire.

Table 1 relates each of the known J_H region amino acid sequences to the J_H gene that apparently encodes it. To account for differences between the genes and the amino acid sequences, two somatic mechanisms were invoked: truncation of *J* regions during *V–J* recombination (see below), indicated in Table 1 by an X, and base substitutions, for which a different amino acid residue is indicated. There are six cases in which base substitution may have occurred; however, the Val/Trp interchange in Hdex1 may reflect a different allele of the J_{H1} gene because Hdex1 is derived from the (BAB-14 × BRVR)F_1 strain (18). Moreover, the apparent MOPC21 interchange is Asp/Asn, and these are often difficult to distinguish in amino acid sequencing. There are thus only four clear examples of J_H sequences that have suffered somatic mutations: W3207, U61, HOPC1, and MOPC173. Therefore, of 458 sequenced J_H amino acid residues, only four changes have resulted from somatic mutation.

Junctional Variation During *V–J* Recombination. The J_H region expressed in HPC76 (Fig. 2) is derived from the germ-line J_{H2} sequence (13), but the first two *J* codons have been deleted during *V–J* recombination. Such truncation, which markedly reduces the length of HV3, occurs frequently. In Table 1, ap-

Table 1. Derivation of J region amino acid sequences from the four J_H genes

Protein	D Segment	J Segment
		J_{H1}
		Y W Y F D V W G A G T T V T V S S
S107,T15, S63, Y5236	YYGSS	- - - - - - - - - - - - - - - - -
H8	YYGNS	- - - - - - - - - - - - - - - - -
M511	GDYGSS	- - - - - - - - - - - - - - - - -
J558	R	- - - - - - - - - - - - - - - - -
Hdex9	R	- - - - - - - - - - - - - - - - -
Hdex2	N	- - - - - - - - - - - - - - - - -
M603	YYGST	X - - - - - - - - - - - - - - - -
M104E	YD	X - - - - - - - - - - - - - - - -
Hdex8	YD	X - - - - - - - - - - - - - - - -
Hdex10	VN	X - - - - - - - - - - - - - - - -
Hdex6	SH	X - - - - - - - - - - - - - - - -
Hdex3	RD	X - - - - - - - - - - - - - - - -
Hdex7	AD	X - - - - - - - - - - - - - - - -
M167	ADYGNSYFG	X X - - - - - - - - - - - - - -
T601	GYYG	X X - - - - - - - - - - - - - -
W3207	YYKYDL	X - - V - - - - - - - - - - - -
Hdex1	NYH	X X - - - - V - - - - - - - - -
		J_{H2}
		Y F D Y W G Q G T T L T V S S
M315	NDHL	- - - - - - - - - - - - - - - -
X24	GYYG	- - - - - - - - - - - - - - - -
Hdex4	KD	- - - - - - - - - - - - - - - -
Hdex5	SNY	- - - - - - - - - - - - - - - -
H76	GVP	X X - - - - - - - - - - - - - -
		J_{H3}
		W F A Y W G Q G T L V T V S A
A4, E109, A47N		X - - - - - - - - - - - - - - -
X44	HYYGYA	X X - - - - - - - - - - - - - -
J539	HYYGYN	X X - - - - - - - - - - - - - -
U61		X - - - - - - - - - - - p - - -
		J_{H4}
		Y Y A M D Y W G Q G T S V T V S S
M173	P	- - - - N - - - - - - - - - - -
H1	GEPPY	- - - - N - - - - - - - - - - -
		(N)
M21	GNYPW	X - - - D - - - - - - - - - - -

The known J_H amino acid sequences of 32 mouse heavy chains (18) are related to the four J_H genes using the minimal number of genetic events. Proposed codon deletions are indicated as X. Amino acid substitutions shown are proposed to result from somatic mutations. Due to incomplete amino acid sequence data, the COOH terminus of some J regions is not shown. Those J region amino acids that form part of the third hypervariable region are shown in **boldface.**

parent codon deletions are indicated by an X. Only 9 V_H sequences include the entire J_H sequence, whereas 12 apparently lack the first codon and 6 lack the first two codons. The absence of the first amino acid encoded by a J_H gene may result from actual deletion of the first J codon, or from recombination within it, in which the first one or two bases of the last V (or D) codon combine with the last two or one base(s) of the first J codon to produce a hybrid codon. In the six cases in which the first two J codons are absent, it is clear that at least the first codon must have been deleted by the recombination event, and for Hdex1

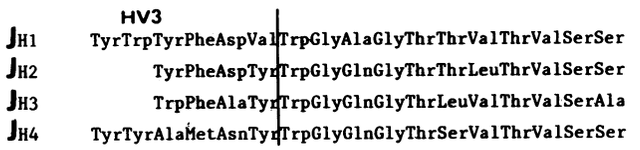

FIG. 4. Comparison of the amino acid sequences encoded by the four authentic J_H genes. The vertical line indicates the border of the third hypervariable region.‡

and HPC76 it is clear that the first two J codons must have been excised.

Pseudo-J_H Gene Segment. We were surprised that the distance between J_{H3} and J_{H4} was as large as 520 base pairs (bp) because the distance between the other J_H genes is only 270 and 336 bp (ref. 13; Fig. 2) and the distance between J_κ genes is conserved [246–310 bp (8 and 9)]. However, midway between J_{H3} and J_{H4} there is a sequence ($J_{H\Psi}$ in Fig. 2) that bears strong homology with the other J_H genes and may represent a pseudo-J_H gene. Fig. 5 compares the sequences of the four genuine J_H genes and their 5′ flanking regions with the equivalent sequence for $J_{H\Psi}$. The authentic J_H sequences have been aligned to maximize homology, which entails introducing deletions within their 5′-flanking sequences. To align $J_{H\Psi}$ required three insertions (of 1, 2, and 5 bp) within the "coding region" and an insertion (1 bp) and a deletion (8 bp) within the 5′-flanking sequence. The nucleotides that are boxed represent the "consensus" sequence at the position—i.e., the nucleotide that occurs in all, or all but one, J_H gene. At numerous positions, there is no consensus sequence—i.e., the genuine J_H genes are themselves highly diverged. However, there are four regions (A–D in Fig. 5) where there are blocks of conserved sequence. Within these regions, $J_{H\Psi}$ differs from the consensus sequence by little more than does any other J_H gene. $J_{H\Psi}$ may be analogous to $J_{\kappa 3}$, which does not correspond to any published κ sequence (8, 9). It is problematical whether a V_H gene could recombine with $J_{H\Psi}$. Although the putative recombination recognition sites 5′ to the J_{H4} coding sequence (A and B in Fig. 5) are homologous with those of other J_H genes, we do not know whether recombination is possible because the distance between them (15 bp) is not the conserved 22 bp found for other J_H genes. Even if a V_H gene were to recombine with this $J_{H\Psi}$ segment, mRNAs containing this J would not encode full-length polypeptides because all the reading frames of $J_{H\Psi}$ contain in-phase termination codons.

V–J Recombination and the Stochastic Model of Allelic Exclusion. Expression of immunoglobulin heavy chain genes is confined to one allele. Allelic exclusion may be explained by a stochastic model involving an inefficient or error-prone recombination machinery (25). Incorrect V–J joining could occur in two ways. First, recombination could put the J gene out of phase with the V gene. Certainly it is clear that a V (or D) region can recombine with a J_H gene at several positions. Unless there is some mechanism to prevent out-of-phase recombination, as many as two of three V–J recombinations could be nonproductive. Second, a V gene could recombine with a pseudo-J sequence rather than with the J gene proper. Although we do not know if V_H–$J_{H\Psi}$ recombination is possible, there is an example of such an "abortive" translocation in the λ light chain system. In the plasmacytoma J558, both copies of the $V_{\lambda 1}$ gene are rearranged, but conventional $V_{\lambda 1}$–$J_{\lambda 1}$ recombination has taken place for only one allele, whereas the other rearranged $V_{\lambda 1}$ gene has recombined with a J_λ-like nucleotide sequence which contains in-phase termination codons (unpublished data). Thus, it is clear that aberrant recombination is one of the mechanisms of allelic exclusion.

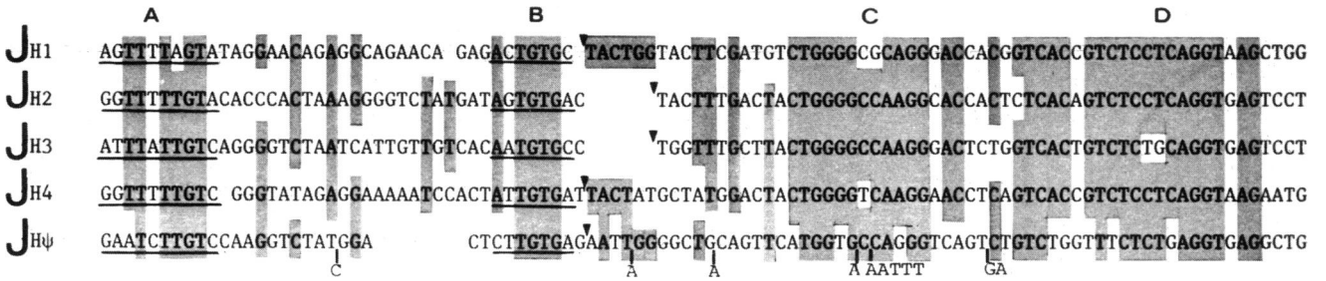

FIG. 5. Comparison of nucleotide sequences of the four authentic J_H genes and $J_{H\Psi}$. The regions that are identical in all (or all but one) J regions are boxed. In order to align $J_{H\Psi}$ with the other J_H genes, deletions and insertions were allowed; the deleted nucleotides are shown below the $J_{H\Psi}$ sequence, and a short line indicates their position. The first coding nucleotide of each J_H gene is indicated by an arrowhead. The common sequences 5′ to J_H genes are underlined. A, B, C, and D indicate regions of maximal homology.

We thank Alison Dredge for excellent technical assistance and Jerry Adams and Suzanne Cory for their critical reading of the manuscript. This work was supported by the National Cancer Institute (R01 CA12421), the American Heart Association, and the National Health and Medical Research Council (Canberra).

1. Brack, C., Hirama, M., Lenhard-Schüller, R. & Tonegawa, S. (1978) *Cell* **15**, 1–14.
2. Lenhard-Schüller, R., Hohn, B., Brack, C., Hirama, M. & Tonegawa, S. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4709–4713.
3. Seidman, J. G. & Leder, P. (1978) *Nature (London)* **276**, 790–795.
4. Davis, M. M., Calame, K., Early, P. W., Livant, D. L., Joho, R., Weissman, I. L. & Hood, L. (1980) *Nature (London)* **283**, 733–739.
5. Cory, S., Adams, J. & Kemp, D. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4943–4947.
6. Bernard, O., Hozumi, N. & Tonegawa, S. (1978) *Cell* **15**, 1133–1144.
7. Seidman, J. G., Max, E. E. & Leder, P. (1979) *Nature (London)* **280**, 370–375.
8. Max, E. E., Seidman, J. G. & Leder, P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3450–3454.
9. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. (1979) *Nature (London)* **280**, 288–294.
10. Gilmore-Hebert, M. & Wall, R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 342–345.
11. Schibler, U., Marcu, K. B. & Perry, R. P. (1978) *Cell* **15**, 1495–1509.
12. Weigert, M., Perry, R., Kelly, D., Hunkapiller, T., Schilling, J. & Hood, L. (1980) *Nature (London)* **283**, 497–499.
13. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. (1980) *Cell* **19**, 981–992.
14. Bernard, O. & Gough, N. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3630–3634.
15. Gough, N. M., Kemp, D. J., Tyler, B. M., Adams, J. M. & Cory, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 554–558.
16. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **75**, 560–564.
17. Vrana, M., Rudikoff, S. & Potter, M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1957–1961.
18. Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. (1980) *Nature (London)* **283**, 35–40.
19. Rocca-Serra, J., Milili, M. & Fougereau, M. (1975) *Eur. J. Biochem.* **59**, 511–523.
20. Adams, J. M., Gough, N. M., Webb, E. A., Tyler, B. M., Jackson, J. & Cory, S. (1980) *Biochemistry* **19**, 2711–2719.
21. Cory, S., Jackson, J. & Adams, J. (1980) *Nature (London)* **285**, 450–455.
22. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. (1980) *Nature (London)* **286**, 676–683.
23. Padlan, E. A. (1977) *Q. Rev. Biophys.* **10**, 35–65.
24. Cook, W. D., Dharmgrongartama, B. & Scharff, M. D. (1979) in *The Cells of Immunoglobulin Synthesis*, ed. Pernis, B. & Vogel, H. J. (Academic, New York), pp. 99–112.
25. Adams, J. M. (1980) *Immunol. Today* **1**, 10–17.