DNA Sequences of the Joining Regions of Mouse $\lambda $ Light Chain Immunoglobulin Genes

Bonnie Blomberg; Susumu Tonegawa

Proceedings of the National Academy of Sciences of the United States of America, Vol. 79, No. 2, [Part 1: Biological Sciences] (Jan. 15, 1982), 530-533.

Stable URL:
http://links.jstor.org/sici?sici=0027-8424%2819820115%2979%3A2%3C530%3ADSOTJR%3E2.0.CO%3B2-Y

Proceedings of the National Academy of Sciences of the United States of America is currently published by National Academy of Sciences.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/nas.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Thu Feb 3 14:26:45 2005

Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 530–533, January 1982
Genetics

DNA sequences of the joining regions of mouse λ light chain immunoglobulin genes

(V-J joining/pseudo λ chain/λ immunoglobulin chain evolution)

BONNIE BLOMBERG* AND SUSUMU TONEGAWA*

Basel Institute for Immunology, 487 Grenzacherstrasse, Basel, CH-4005, Switzerland

Communicated by H. N. Eisen, September 14, 1981

ABSTRACT The joining (J) segments of mouse immunoglobulin λ light chain genes, λ2, λ3, and a presumptive λ4, were cloned, and their sequences were determined and compared with that of λ1. Although all the λ J segments share sequence homology, the J1 and J4 segments and the J2 and J3 segments, respectively, are more homologous. These sequence data, together with the fact that present day λ genes occur in two clusters, 5', $J_3C_3J_1C_1$ 3' and 5' $J_2C_2J_4C_4$3', further substantiates a probable evolutionary duplication unit, $J_{II}C_{II}J_IC_I$, with II the precursor of λ3 and λ2 and I the precursor of λ1 and λ4. From the J4 sequence, we conclude that the λ4 gene is most likely nonfunctional (i.e., a pseudogene). The signal nonamer sequence 5' to J3 differs from that of J1 in two consecutive base pairs. This difference could account in part for the lower level of expression of λ3 as compared with λ1 in mouse serum.

A mammalian immunoglobulin molecule is composed of one of two types of light chain, κ or λ, in association with a heavy chain. The percentage of serum κ and λ immunoglobulin varies in different species: human immunoglobulin is ≈60% κ and 40% λ (1) whereas in mouse immunoglobulin is 95% κ and 5% λ (2, 3). In mouse, three λ chain subtypes are known: λ1(4–6), λ2(7), and λ3(8), which occur in the serum in the approximate ratio 8:1:1 (3, 8). A light chain is encoded by three DNA segments, variable (V), joining (J), and constant (C), that are separate in the germ line (9, 10). A subtype is defined by its C region amino acid sequence. There are only two V region genes, V1 and V2, for mouse λ(9). All λ1 chains studied to date contain V regions encoded by V1 (9, 11). Likewise, all known λ2 chains contain V regions encoded by V2 (3, 8, 12). It has recently been shown that λ3 chains also use the V1 gene (13, 14). We also discovered a fourth Cλ gene, C4 (13), by virtue of its crosshybridization with Cλ1; however, no λ chain corresponding to C4 has been described.

It has been shown that mouse λ light chain C genes occur in two clusters: $J_3C_3J_1C_1$ (13, 15) and $J_2C_2J_4C_4$ (13). Since the λ1 and λ2 V and C genes have been mapped to chromosome 16 (16), the λ3 and λ4 loci must also be located on this chromosome. Previous studies showed that the κ light chain J DNAs (17, 18), as well as heavy chain J DNAs (19), are closely linked and have no C gene segments between the J DNA segments. Therefore the organization of the J and C DNA segments in the mouse λ system demonstrates an alternative evolutionary duplication unit for immunoglobulin J and C gene segments: in the case of κ and heavy chains, the J segments duplicated separately from their corresponding C segments whereas, with λ chains, a J and C segment duplicated as a unit.

We determined the DNA sequences of the λ J segments of J2, J3, and J4 for three reasons: to obtain further evidence for an evolutionary duplication unit in λ, to determine whether the J_4C_4 gene segments could be functional, and to seek a molecular basis for the differential expression of λ subtypes.

MATERIALS AND METHODS

Bacteria and Phages. Phage Charon 4A was obtained from F. Blattner (Univ. of Wisconsin, Madison, WI) (20). The λgt WES-λB was obtained from P. Leder (National Institutes of Health) (21). *Escherichia coli* 803 (r_k^-, m_k^-, Su III$^+$) was originally from K. and N. E. Murray (University of Edinburgh). Lysogens used for preparation of packaging mixtures, BHB 2688 [N205 recA$^-$ (λ imm$_{434}$ b2*red3 E*am4 Sam7)/λ], and BHB 2690 [N205 recA$^-$ (λ imm$_{434}$ *cI*$_{ts}$ b2*red3 E*am15 Sam7)/λ] were obtained from B. Hohn (Friedrich Miescher Institute, Basel) (22).

Preparation of Libraries and Isolation of Clones. Preparation of the mouse embryo (BALB/c) *Eco*RI partial library has been described (23). Clones IgS8.2, Ig10A1, and Ig16.E5 were isolated with either the Cλ1 or the (V + C)λ2 probe as described (13). The preparative agarose gel technique has been described (24) and DNA was prepared as before (25). The procedure to prepare nick-translated probes (9) and the screening of the recombinant phage carrying mouse DNA (26) have also been described. Restriction enzymes were purchased from Boehringer Mannheim, New England BioLabs, or Bethesda Research Laboratories.

Nucleotide Sequence Analysis. Sequence analysis was performed according to the method of Maxam and Gilbert (27). Products of analysis reactions were loaded once onto a 20% and three times sequentially onto an 8% acrylamide gel (1.5 × 200 × 400 mm) containing 7 M urea/100 mM Tris borate, pH 8.3/2 mM EDTA.

RESULTS AND DISCUSSION

Sequence Analysis Strategy for the DNA Segments of λ J2, J3, and J4. The cloned DNA fragments carrying the four mouse λ light chain J and C gene segments have been described (13, 28). The organization of these gene segments in the germ-line genome is summarized in Fig. 1. Each C gene segment carries its own J segment ≈1.3 kilobases (kb) upstream (to its 5' side) (13). The nucleotide sequence of J1 has been determined (28).

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: bp, base pair(s), C, J, and V, constant, joining, and variable regions, respectively; kb, kilobase(s).

* Present address: Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139.

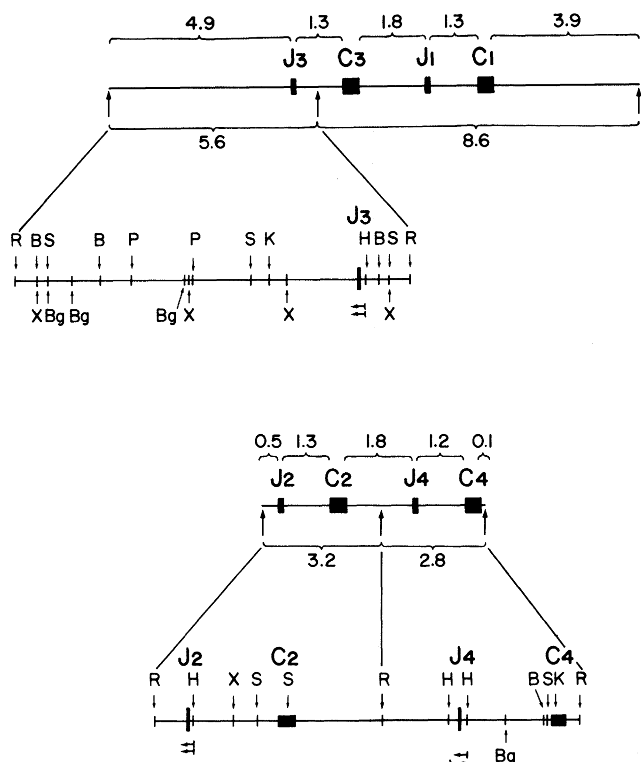

FIG. 1. Germ-line configuration of mouse λ J and C segments and restriction endonuclease cleavage maps of *Eco*RI inserts containing J2, J3, and J4. The $J_3C_3J_1C_1$ gene segments are from clone IgS8.2 and the $J_2C_2J_4C_4$ segments are from clone Ig10A1 (ref. 13), which contained DNA from *Eco*RI partial digests of BALB/c embryo DNA. Distances are in kb. Large arrows designate *Eco*RI cleavage sites. The distance between the two clusters is unknown. Each *Eco*RI insert was purified by electrophoresis in acrylamide and elution. Insert DNA was then digested with one or more enzymes. Horizontal arrows indicate direction and extent of sequence determination. B, *Bam*HI; Bg, *Bgl* II; H, *Hin*dIII; K, *Kpn* I; P, *Pst* I; R, *Eco*RI; S, *Sac* I; X, *Xba* I.

To determine the nucleotide sequences of the J2, J3, and J4 segments, we constructed restriction enzyme maps of the *Eco*RI inserts carrying these DNA segments (Fig. 1). Since *Hin*dIII cut very close [within 100 base pairs (bp)] to the 3′ side of each J segment, the appropriately labeled *Hin*dIII fragments were prepared and their sequences were analyzed (Fig. 2).

The $J_3C_3J_1C_1$ and $J_2C_2J_4C_4$ Sequences Arose by Duplication. The DNA sequences of the four presumptive J segments and their surrounding regions are compared in Fig. 2. In the case of J1, J2, and J3, the DNA sequences correspond to the previously determined amino acid sequences (4, 6, 14, †). These results confirm the existence of a J segment to the 5′ side of each of the three C segments, C1, C2, and C3. For J4, no amino acid sequence is available; however, a J-like sequence exists ≈1.3 kb 5′ to the C4 segment. This confirms our previous R-loop analysis for the presence of the J4 segment (13).

The J1 and J4 sequences and the J2 and J3 sequences, respectively, are clearly more alike, especially in the noncoding regions. The percent homologies of the four λ J DNAs for the areas within the protein-encoding regions of the surrounding regions are compared in Fig. 3. It appears that the most recent evolutionary duplication involved a precursor $J_{II}C_{II}J_IC_I$ giving rise to the present day $J_3C_3J_1C_1$ and $J_2C_2J_4C_4$. The areas outside the J-encoding regions are similar for the J1-J4 and J3-J2 pairs, 74% and 83% homology, respectively. These areas are much less similar (25–39%) for the other four comparisons. Thus, the duplication of $J_{II}C_{II}J_IC_I$ occurred much more recently than that of a precursor for $J_{II}C_{II}$ and J_IC_I.

Fig. 3 also shows that the areas outside the J segments have diverged more than the coding regions. This is especially evident for comparisons of the more evolutionarily distant J segments (i.e., J1-J2, J1-J3, J4-J3, and J4-J2). This argues for an evolutionary selection for conservation of the J region sequences. The J amino acid sequence is highly conserved, not only among the λs but also for all κ and λ light chains (see refs. 1 and 17 and below). The J region immediately adjoins the third hypervariable or complementarity-determining region and may be used to position this region of light and heavy chains to allow better contact with the antigenic determinant (29).

The λ4 Gene is Probably a Pseudogene. In Fig. 2, the J4 sequence is presented to give maximal homology to the other three λ J sequences. It appears that J4 has undergone a 2-bp deletion in the signal heptamer and a 1-bp insertion (thymine) immediately after the codon for amino acid position 100. Because of this insertion, the amino acid sequence encoded by J4 in the same reading frame as the first tripeptide, Trp-Val-Phe, is very different from those of the other light chain J sequences, all of which have glycine in positions 101 and 103 (1). In the second frame, a termination codon (TGA) interrupts coding while, in the third frame, the amino acid encoded at position 99 (by TGT) is cysteine, which does not appear in any of the J regions studied to date. Furthermore, in all functional J segments, a dinucleotide G-T, which is an obligatory part of an RNA splicing signal (30), occurs at the position corresponding to amino acid residue 110 (Fig. 3 and ref. 17). In contrast, this dinucleotide is absent in all reading frames of J4 at this position. The absence of a proper splice signal was previously observed in Jκ3 (17, 18). In addition, no protein representative of λ4 has been described and all serum, hybridoma, and myeloma λ chains can be accounted for by λ1, λ2, and λ3 (3, 4, 8). We therefore believe that J4 does not participate in the synthesis of a functional λ chain and thus that λ4 may be a pseudogene.

Residue 98 May Be Encoded by the V Side or the J Side or Both. The 3′ ends of germ-line Vλ and the 5′ ends of Jλ DNA segments are shown in Fig. 4. For λ1, as has been noted (28), the codon TGG for tryptophan at position 98 can be generated in one of two ways: the first base (thymine) can be provided by the V region and the rest by the J region or all three bases can be provided by the J region. For the λ3 protein, where V1 is joined with J3, the recombination sites may not be defined so precisely. Position 98 for λ3 is phenylalanine, which may be generated totally or partially from either the V or the J side (Fig. 4). Misphasing may occur more easily in this area of repeated thymine: we have noticed an example of a joined V1-J3 in myeloma J558 that contains a thymine deletion in the codon for residue 98 (unpublished). The first two or all three bases of the phenylalanine codon at position 98 of the λ2 chain arise from Vλ2.

The J3 Signal Sequence Differs from that of J1. The detailed mechanism of V-J joining is unknown, but deletion of the germ-line DNA sequences between the V and J regions has been shown (17, 31). The joining most likely involves recognition of conserved nonamer and heptamer signal sequences 3′ to the V region and 5′ to the J region by a putative "recombinase" (19, 32). Base pairing between these sequences may enhance the fidelity of V-J joining (Fig. 4) but by showing these stem structures we do not wish to imply the mechanism of recombination.

Although λ1 and λ3 can share the same V gene, the level of expression of λ3 is much lower than that of λ1. In the serum, the combined level of λ3 and λ2 is ≈20% of the total λ chains

† Azuma, T., Steiner, L. A. & Eisen, H. N. (1980) Fourth International Congress of Immunology, Paris, 1.1.01 (abstr.).

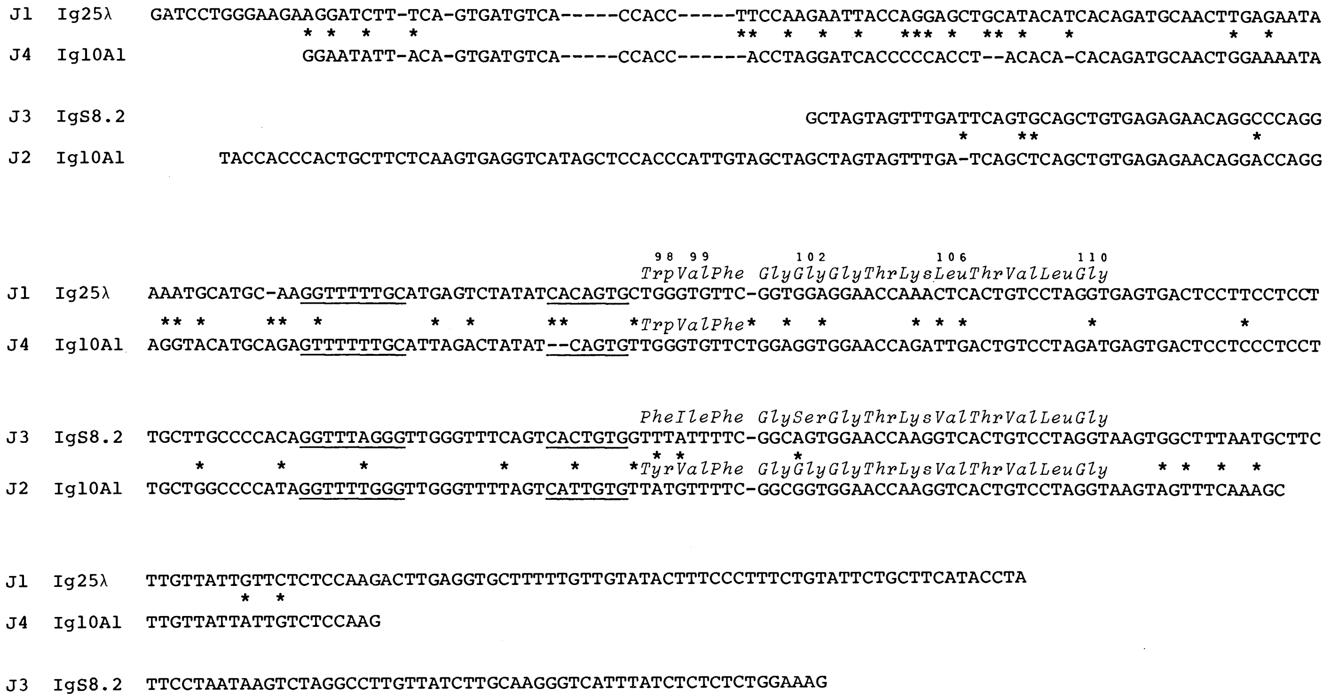

FIG. 2. Comparison of nucleotide sequences of germ-line λ J segments and surrounding regions. *, Nonidentical base pairs in comparisons of the J1-J4 and J3-J2 sequences. Signal nonamer and heptamer sequences 5′ to the J regions are underlined. Amino acids encoded by the nucleotide sequences are shown in italics. The J1 DNA sequence is taken from ref. 28. The J2 fragment for sequence analysis was prepared from phage DNA from clone Igl0A1. The DNA was digested with *Hin*dIII and the 3′ ends were labeled with [α-^{32}P]NTPs using the Klenow fragment of *Escherichia coli* DNA polymerase I as described (17). After further cleavage with *Eco*RI, the 500-bp fragment was isolated on and eluted from 5% acrylamide (procedure 9 of ref. 27). The J3 fragment was prepared in a similar fashion except that 2 μg of the purified 5.5-kb *Eco*RI insert from clone IgS8.2 was digested with *Hin*dIII, end labeled, and then digested with *Xba* I. The 1100-bp fragment was eluted from acrylamide and its sequence was analyzed. The J4 fragment was prepared similarly: Igl0A1 phage DNA or insert DNA (2.8 kb) was digested with *Hin*dIII and end labeled. The 270-bp fragment was isolated on acrylamide and strand separated (procedure 8 of ref. 27) and its sequence was analyzed. The AAG triplet at the 3′ end of each of these sequences is most likely part of the *Hin*dIII site (A-A-G-C-T-T).

in immunoglobulin. From the ratio of λ3 to λ2 hybridomas (R. Zaugg, V. Igras, and H. N. Eisen, personal communication), we assume that approximately half of this λ3/λ2 is λ3 and therefore, in the serum, λ3 is probably <10–15% of the λ1 level. In addition, it has recently been found that the biosynthesis of λ

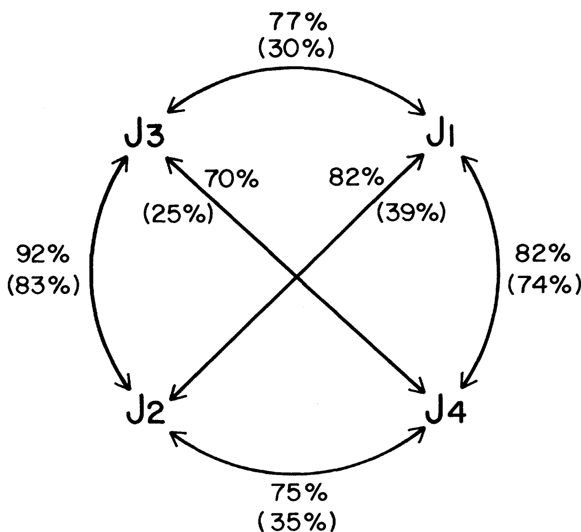

FIG. 3. Percent homologies of DNA sequences of the λ J segments. Values for the protein-encoding regions (39 or 40 bp for each J segment) are given for each pair compared. Homologies for the areas outside the J segments (parentheses; in bp) for the total comparable sequences are J1-J4, 115/155; J3-J2, 84/101; J1-J3, 49/163; J1-J2, 59/152; J4-J3, 31/124; J4-J2, 50/144.

chain by spleen cells also favors λ1 over λ3 with the level of λ3 about one third that of λ1 (E. B. Reilly, A. R. Frackelton, Jr., and H. N. Eisen, personal communication). There are two consecutive base pairs changed within the nonamer signal sequence of J3 as compared with J1 (Figs. 2 and 4), which may produce less efficient V1-J3 joining and hence a lower level of λ3 expression. We cannot rule out the possibility that amino acid differences in the J regions may contribute to differential antigen binding and hence antigen selection, but we feel that the differential antigen selection would be minimal because the same V1 is used for both light chains. Therefore, the level of expression of λ1 and λ3 may reflect the efficiency of V-J joining.

This proposed V-J joining regulation mechanism may not apply to the differential expression of λ1 and λ2. Although the level of λ2 is lower in the serum, the cytoplasmic λ1 and λ2 are approximately equal (E. B. Reilly, A. R. Frackelton, Jr., and H. N. Eisen, personal communication). Here the effect of antigenic selection may be more pronounced: λ1 and λ2 use different V regions. Alternatively, λ1 may be amplified by a cellular mechanism involving subtype recognition. The signal sequences of λ2 do not differ greatly from those of λ1 or the other consensus sequences and V-J joining may be of similar efficiency for λ1 and λ2.

There are no direct data to indicate how changes in these signal sequences might influence expression of a given immunoglobulin light or heavy chain, but some preferential DNA joinings have been noted in diversity-J joining (D-J) in the immunoglobulin heavy chain. JH_2 may be preferred and JH_3 is rarely used (unpublished results). The JH_2 sequence has the consensus nonamer sequence and JH_3 differs in three of the nine bases (19).

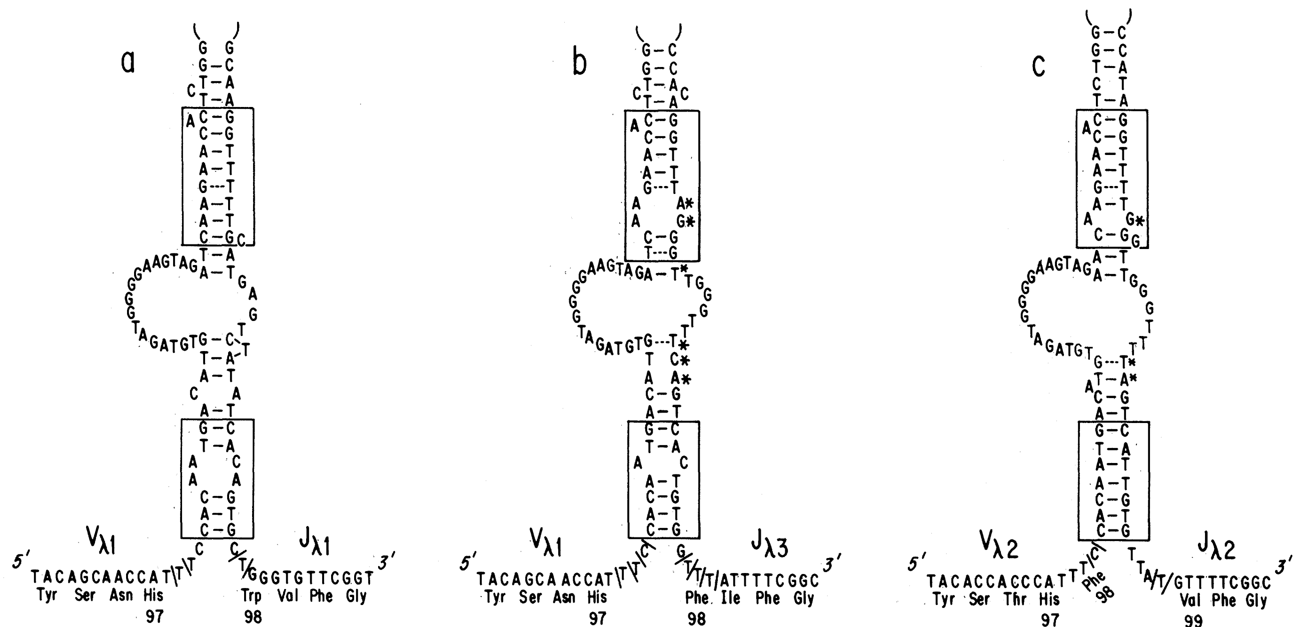

FIG. 4. Inverted repeat stem structures formed between the 3′ noncoding regions of embryonic Vλ1 and Vλ2 DNAs and 5′-flanking regions of their corresponding J DNAs. (*a*) Vλ1 and Jλ1. (*b*) Vλ1 and Jλ3. (*c*) Vλ2 and Jλ2. The sequences for the antisense strand were omitted. The conserved nonamer and heptamer sequences are boxed. Oblique lines indicate possible recombination sites. *, Bases different from the Jλ1 sequence.

We thank Dr. Herman N. Eisen for allowing us to refer to his unpublished results. The Basel Institute for Immunology was founded and is supported by Hoffmann–La Roche.

1. Kabat, E. (1976) *Structural Concepts in Immunology and Immunochemistry* (Holt, Rinehart and Winston, New York), p. 232.
2. McIntire, K. R. & Rouse, A. M. (1970) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **29**, 704 (abstr.).
3. Cotner, T. & Eisen, H. N. (1978) *J. Exp. Med.* **148**, 1388–1399.
4. Weigert, M., Cesari, I. M., Yonkovich, S. J. & Cohn, M. (1970) *Nature (London)* **228**, 1045–1047.
5. Appela, E. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 590–594.
6. Cesari, I. M. & Weigert, M. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2112–2116.
7. Dugan, E. S., Bradshaw, R. A., Simms, E. S. & Eisen, H. N. (1973) *Biochemistry* **12**, 5400–5416.
8. Azuma, T., Steiner, L. A. & Eisen, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 569–573.
9. Brack, C., Hirama, M., Lenhard-Schuller, R. & Tonegawa, S. (1978) *Cell* **15**, 1–14.
10. Hozumi, N. & Tonegawa, S. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3628–3632.
11. Weigert, M. & Riblet, R. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 837–846.
12. Elliot, B. W., Jr., Steiner, L. A. & Eisen, H. N. (1981) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **40**, 1098 (abstr.).
13. Blomberg, B., Traunecker, A., Eisen, H. N. & Tonegawa, S. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3765–3769.
14. Breyer, R., Sauer, R. & Eisen, H. N. in *Immunoglobulin Idiotypes*, ICN-UCLA Symposia on Molecular and Cellular Biology, eds. Janeway, C. A., Sercarz, E. E., Wigzell, H. & Fox, C. F. (Academic, New York), in press.
15. Miller, J., Bothwell, A. & Storb, U. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3829–3833.
16. D'Eustachio, P., Bothwell, A. L. M., Takaro, J. K., Baltimore, D. & Ruddle, F. H. (1981) *J. Exp. Med.* **153**, 793–800.
17. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. (1979) *Nature (London)* **280**, 288–294.
18. Max, E. E., Seidman, J. F. & Leder, P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3450–3454.
19. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. (1980) *Nature (London)* **286**, 676–683.
20. Blattner, F. R., Williams, G. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977) *Science* **196**, 161–169.
21. Leder, P., Tiemeier, D. & Endquist, L. (1977) *Science* **196**, 175–177.
22. Collins, J. & Hohn B. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4242–4246.
23. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2138–2142.
24. Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. & Tonegawa, S. (1979) *Nature (London)* **277**, 627–633.
25. Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wabl, M., Raschke, W. & Tonegawa, S. (1981) *Cell* **24**, 353–365.
26. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
27. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
28. Bernard, O., Hozumi, N. & Tonegawa, S. (1978) *Cell* **15**, 1133–1144.
29. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1979) *Sequences of Immunoglobulin Chains* (NIH Publication no. 80-2008).
30. Breathnach, R., Benoist, C., O'Hare, K., Cannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
31. Seidman, J. G., Nau, M. M., Norman, B., Kwan, S.-P., Scharff, M. & Leder, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6022–6026.
32. Early, P., Huang, H., Davis, M. M., Calame, K. & Hood, L. (1980) *Cell* **19**, 981–992.