Proc. Natl. Acad. Sci. USA
Vol. 75, No. 5, pp. 2429–2433, May 1978
Immunology

Variable region genes for the immunoglobulin framework are assembled from small segments of DNA—A hypothesis

(minigenes/insertional mechanism)

ELVIN A. KABAT*, TAI TE WU†, AND HOWARD BILOFSKY‡

* The National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014; and Departments of Microbiology, Human Genetics and Development, and Neurology, College of Physicians and Surgeons, Columbia University, New York, New York 10032; † Departments of Engineering Science and Biochemistry and Molecular Biology, Northwestern University, Evanston, Illinois 60201; and ‡ Bolt, Beranek and Newman, Inc., Cambridge, Massachusetts 02138

Contributed by Elvin A. Kabat, March 6, 1978

ABSTRACT **Sequences of each of the four framework segments FR1, FR2, FR3, and FR4 of the variable regions (V-regions) of light and heavy chains of immunoglobulins were grouped into sets with identical sequences. Sets contained from 1 to 18 members. When each V-region was traced from one FR to the next, it was seen that members of the same set in FR1 could be associated with different sets in FR2, FR3, and FR4. This suggests that the framework for the light and heavy chain V-regions is assembled during embryonic development from sets of minigenes for each FR segment. FR4 from three sets of human V_κI chains also contained members of V_κII, V_κIII, and V_κIV subgroups; one FR2 set contained eight rabbit V_κ, one human V_κIV, and four mouse V_κ and an FR4 set contained two human V_HIII and one mouse V_HIII, indicating substantial evolutionary preservation of these sequences and suggesting that the sets of minigenes are highly conserved in the germ line. The clone of Tonegawa *et al.* [Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* 75, 1485–1489] could be a hybrid FR1 and FR3 coming from minigenes of MOPC 315 (a V_λII) whereas FR2 would come from MOPC 104E (a V_λI). That FR4 is not joined to the rest of the V-region in 12-day-old mouse embryo DNA is also in accord with this hypothesis. Mouse sperm DNA should be examined to establish whether the hypothesized minigenes are separated by intervening sequences and whether the complementarity-determining (hypervariable) regions or segments of the V-region are separated from the framework in genomic DNA. Sperm DNA from rabbits or other species could be used to search for minigene segments whose sequences are identical in several species.**

The variable region (V-region) of immunoglobulin light and heavy chains may be divided into four framework (FR) and three complementarity-determining [hypervariable (1)] regions or segments (CDR). In the V-region of light chains (V_L), FR1, FR2, FR3, and FR4 comprise residues 1–23, 35–49, 57–88, and 98–107, respectively, and CDR1, CDR2, and CDR3 comprise residues 24–34, 50–56, and 89–97, respectively (1, 2). CDR1 and CDR3 may vary in length due to insertions (1–3, cf. 4). In the V-region of heavy chains (V_H), the FR are 1–30, 36–49, 66–94, and 103–113 and CDR1, CDR2, and CDR3 are residues 31–35B, 50–65, and 95–102; CDR1, CDR2, CDR3, and FR3 may also have insertions (cf. 2). The light (5–7) and the heavy chains (8, 9) have been divided into subgroups based essentially on sequences in FR1. Each framework sequence for a subgroup is considered to represent a single structural gene. Recently, Potter (10) has classified mouse V_κ sequences into groups called "isotypes," each isotype having sequences in FR1 that could have up to three substitutions hypothesized as resulting from somatic mutation. Efforts to subgroup beyond FR1 have been relatively uninformative.

In the course of preparing the second edition of *Variable Regions of Immunoglobulin Chains* (2), we attempted to use the PROPHET computer to organize the listing of the light and heavy chains of the various species by selecting identical FR1, FR2, FR3, and FR4 segments; these were then arranged so that one could readily see sets identical in FR1 and within such sets one could see identical FR2, FR3, and FR4. It became evident that chains with identical FR1 could be associated with different sets of FR2, FR3, and FR4. These findings are best interpreted on the hypothesis that FR1, FR2, FR3, and FR4 consist of germ-line dictionaries of minigenes and that, in the process of assembly of the complete V-region gene, these are selected by joining the appropriate FR and CDR segments of DNA. The recently described clone of Tonegawa *et al.* (11), for which the nucleic acid of the mouse V_λ gene was sequenced (12) from 12-day-old mouse embryo DNA, can be interpreted as having been assembled from FR1 and FR3 from MOPC 315 (a V_λII) and FR2 (except for Glu vs. Gln as discussed below) from MOPC 104E (a V_λI); CDR1 had the same sequence in both V_λI and V_λII whereas CDR2 and CDR3 had sequences of both V_λI and V_λII. In line with recent findings in viral and mammalian genes, the present analysis predicts that intervening sequences will be found between the FR fragments and that the CDR will be recombined into the FR segments during early embryonic development with elimination of the intervening sequences to assemble a V-region gene. Intervening sequences (13) have now been found in a mouse β-globin (14, 15), rabbit globin (16), avian ovalbumin (17, 18), viral proteins (19–25), some ribosomal DNA genes in *Drosophila* (26–29), and tRNA genes in yeast (30).

RESULTS

The data base (2) of V-region sequences supplemented by additional published data (31–46) contained complete and partial sequences of 509 light and 225 heavy chains. Sequences of human V_κI, mouse V_κ, rabbit V_κ, and human and mouse V_HIII, the only groups with sufficient data, were sorted so that identical sequences of FR1, FR2, FR3, and FR4 were located; Glx and Asx were accepted as equivalent to Gln or Glu and Asn or Asp in assembling sets. These were grouped further and ex-

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

Abbreviations: V-region, variable region; V_L, V_H, variable regions of light and heavy chains, respectively; V_κ, variable regions of κ light chains; V_λI, V_λII, variable regions of two subgroups of mouse λ light chains; V_HIII, subgroup of heavy chain V-regions; FR, framework regions or segments of the V-region; CDR, complementarity-determining (hypervariable) regions or segments of the V-region; C-region, constant region of light and heavy chains.

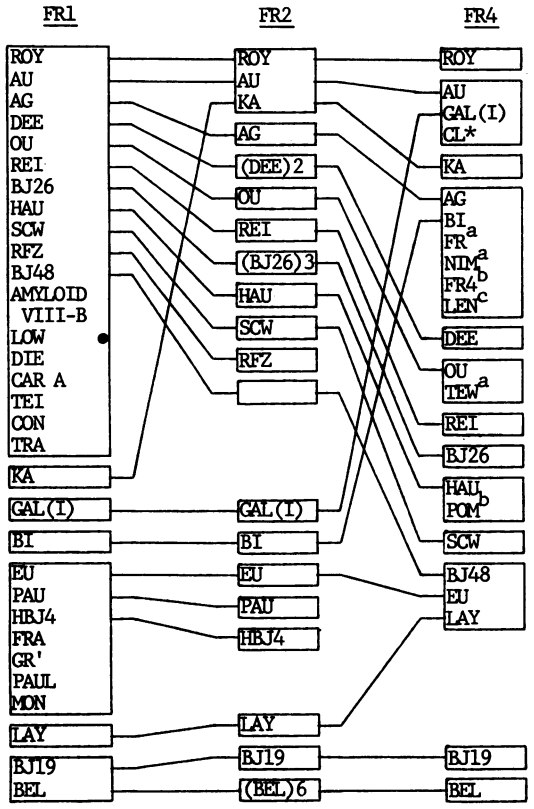

FIG. 1. Assembly of the frameworks of human V_κI chains from minigene sets. ●, Cold agglutinin with anti-blood group I activity; a, human κ light chain subgroup II; b, human κ light chain subgroup III; c, human κ light chain subgroup IV.

amined to determine whether chains in one set in FR1 were associated with one or more sets identical in FR2, FR3, and FR4. Because many V_L and V_H sequences are incomplete and because a sufficient number of sequences is not available, many FR segments had only a single example in a group. This was especially true for FR3. Nevertheless, the data clearly show that chains in a given FR1 set were associated with several sets of FR2, FR3, or FR4. Moreover, members of a given FR set could come from several subgroups or from several species. The findings are given for human V_κI, mouse V_κ, rabbit V_κ, and human and mouse V_HIII. Chains with the same sequence for a given FR are enclosed in a box; a single listing indicates that two identical sequences have not been encountered. If a box is empty, little or no sequence data are available. Sequences that do not permit comparison of two or more FR segments in a given chain have been omitted.

From Fig. 1, on human V_κI chains, it is evident that FR1 contained three sets with identical residues. One set contained 18 chains, another contained 7, and a third 2; there were several other sets with a single sequence. If one follows the lines in Fig. 1, for each chain it is clear that ROY and AU were associated with the same sets of FR1 and FR2 but with different FR4 sets; the third member of the FR2 set, KA, was from FR1 and FR4 sets in which it was the only member. EU, PAU, and HBJ4 from the FR1 with seven proteins each was associated with an individual FR2 set but EU was associated in FR4 with LAY with unique FR1 and FR2. AG from the FR1 set with 18 members and BI from a unique FR1 and FR2 were in the same FR4 set; this FR4 set was unusual in that it contained two V_κII (FR and NIM), one V_κIII (protein FR4), and one V_κIV (LEN); another V_κI (OU) set had an identical V_κII (TEW) and a third (HAU) had an identical V_κIII (POM); all were of human origin. FR3 is not included in Fig. 1 because of incomplete sequences or because no two chains were identical.

Fig. 2 shows the data for mouse V_κ chains. In FR1, one set had eight identical members and two sets had two identical members; in the FR2 set, MOPC 321 and, just recently, TEPC 124, MOPC 63, and Ab22 (D. J. McKean, M. Bell, and M. Potter, personal communication) have been completely sequenced and are identical. Two other proteins could be identical, sequences being incomplete in one and three residues; these are given in parentheses with the number of missing residues alongside. FR3 and FR4 each have two sets with identical members. Again it is evident that members of a given FR1 set may be associated with different sets for FR2, FR3, and FR4.

The data from rabbit V_κ (Fig. 3) are the most striking because there are six sets with more than a single member in FR1 and two in FR2. The top FR2 set with eight rabbit members is associated with two different FR1 sets with three and four members, respectively, plus four each with a single member. The 14 members of FR4 are associated both with unique sets and with sets with several members in both FR1 and FR2 but one member of these sets is associated with a unique FR4 set. The various sets include members with varying antibody ac-

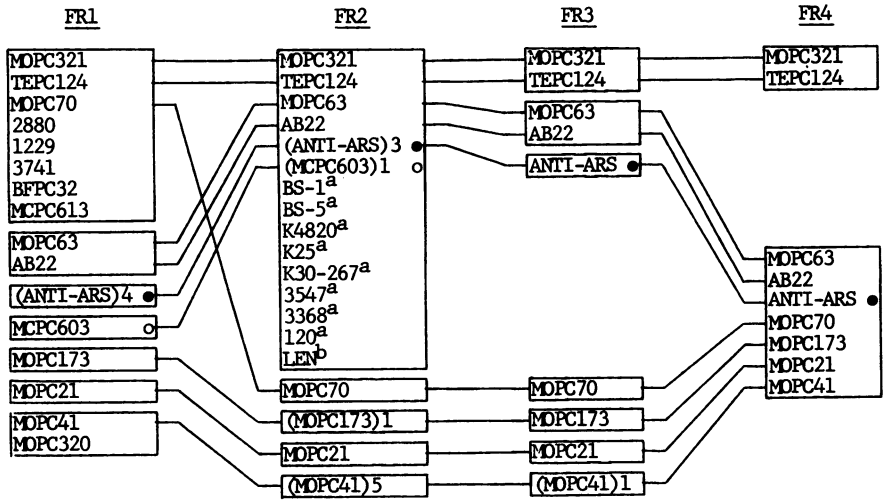

FIG. 2. Assembly of the frameworks of mouse V_κ chains from minigene sets. ●, Anti-*p*-azophenylarsonate; ○, Anti-phosphocholine; a, rabbit κ light chain; b, human κ light chain subgroup IV.

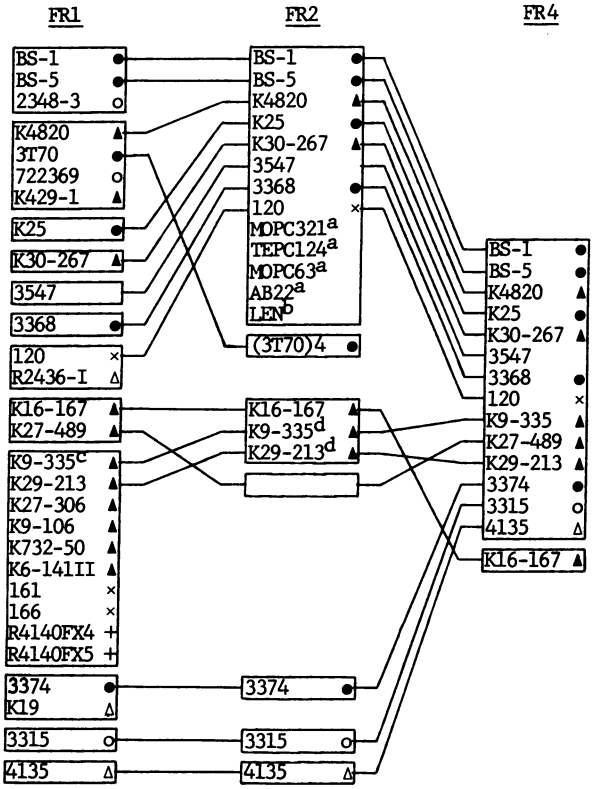

FIG. 3. Assembly of the frameworks of rabbit V_κ chains from minigene sets. ●, Anti-type III pneumococcal polysaccharide; ○, anti-type VIII pneumococcal polysaccharide; ▲, anti-streptococcal group A variant carbohydrate; △, anti-streptococcal group C carbohydrate; +, anti-*p*-azophenylarsonate; ×, anti-*Micrococcus lysodeikticus;* a, mouse κ light chain; b, human κ light chain subgroup IV; c, K9-335 has the same sequence as K9-338; d, the residue at position 36 (2) has been changed from Tyr to Phe for K29-213, K9-335, and K9-338 (35).

tivities; thus, many sets with several members may include antibodies of three different specificities and one with no known antibody activity.

In Figs. 2 and 3 it may be seen that the top FR2 set has eight rabbit, four mouse, and one human V_κIV members. Thus, this sequence of 15 residues has been highly preserved over a substantial evolutionary period. Fig. 4 summarizes combined data for human and mouse V_HIII chains. There are two human and two mouse FR1 sets, three human and two mouse FR2 sets, and two mouse FR3 sets with several members. An FR4 set contains two human and one mouse segment that are identical; the two other groups of mouse FR4 are only potentially identical because their sequences are not complete. It is again evident that members of one set of identical FR1 may be associated with different sets of FR2, FR3, and FR4. The human and mouse sets remain separate except for FR4 in which two human proteins, TIL and NIE, have the same sequence as MOPC 47A, again indicating evolutionary conservation.

When the mouse data are examined for antibody specificity, they fall into two groups—one with phosphocholine specificity and one with $\beta 2 \rightarrow 1$-fructosan specificity. This makes the mouse V_HIII data more limited because they are highly selected, but the human and mouse data agree with the inferences from the data in Figs. 1–3.

DISCUSSION

Despite the relatively small numbers of complete sequences (2, 31–46), the data presented in Figs. 1–4 show a remarkable frequency of occurrence of identical sequences for the individual framework segments, especially those of FR1, FR2, and FR4. The occurrence of identical sequences of FR1 was noted in the original papers in which subgroups of human V_κ chains were established (5–7), and it was pointed out (47) that the occurrence of such high frequency of identical FR1 meant that FR1 could not be involved in complementarity but had to be a structural or framework segment. The data in Figs. 1–4 also make it clear that attempts at subgrouping beyond FR1 could not be successful.

The present finding that members of a given FR1 set can be associated with different FR2, FR3, and FR4 sets is best interpreted on the hypothesis that the corresponding sets of nucleotides in the DNA are separated from one another. Thus, the nucleotide sequence that codes for each framework may be considered to be made up from a set of germ-like minigenes that are assembled during early embryonic development together with the nucleotide sequences for the three CDR to give the complete V-region sequence. This hypothesis is a slight modification of the episomal insertion mechanism proposed earlier (ref. 1; cf. 48) to explain the generation of diversity during embryogenesis by insertion of the nucleotide sequences for the CDR into a framework at a time when the genome was considered to be a linear sequence of individual genes and the only insertional mechanism known was episomal (49). The recent revolution in concepts resulting from the demonstration of leader and intervening sequences makes the insertional hypothesis for the generation of diversity more acceptable. Indeed, the assembly of many genes from pieces (13) is not very different from an insertional mechanism. The recent findings of Tonegawa *et al.* (11, 12) that, in 12-day mouse embryo DNA, the V_λI and C_λI genes are separated by at least 5000 bases whereas in myeloma immunoglobulin they are separated only by 1250 nucleotides indicates that movement of genes and gene segments may be occurring throughout embryonic development. Thus, one must look for the true sequence of a gene in the genome in sperm or in egg DNA because movement of nucleotides may have occurred by the time the embryo is 12 days old.

The conclusions in this study, proposing a set of minigenes for the framework of the V-regions of light and heavy chains, suggest that the findings of Tonegawa *et al.* (12)—that 12-day mouse embryo DNA contains a linear array of nucleotides for residues 1–96 of the V-regions, may represent the result of differentiation and that, in mouse sperm or egg DNA, these nucleotides may not occur in sequence but as minigene segments for each FR region. There might well be intervening sequences between groups of FR minigenes or even between minigenes. The finding that the clone from the mouse embryo DNA that was hybridized yielded a sequence that contained framework residues for V_λII (MOPC 315) and V_λI (MOPC 104E) suggests that such joining had already taken place. That FR4 is not joined to the rest of the V-region sequence suggests that in the cell of the 12-day embryo that donated the DNA sequence the assembly of the V-region had not yet been completed. Mouse sperm DNA constitutes a not too inaccessible source of true genomic DNA and one hopes that it will be used for these studies. Sperm DNA from rabbits or other species could be used to search for those FR segments that are identical in several species. Extreme care must be taken to separate the sperm from the other cells present in semen (50).

This hypothesis will clearly be testable as more sequences become available because the repeats in individual FR segments should increase and some of the groups with a single member should become larger.

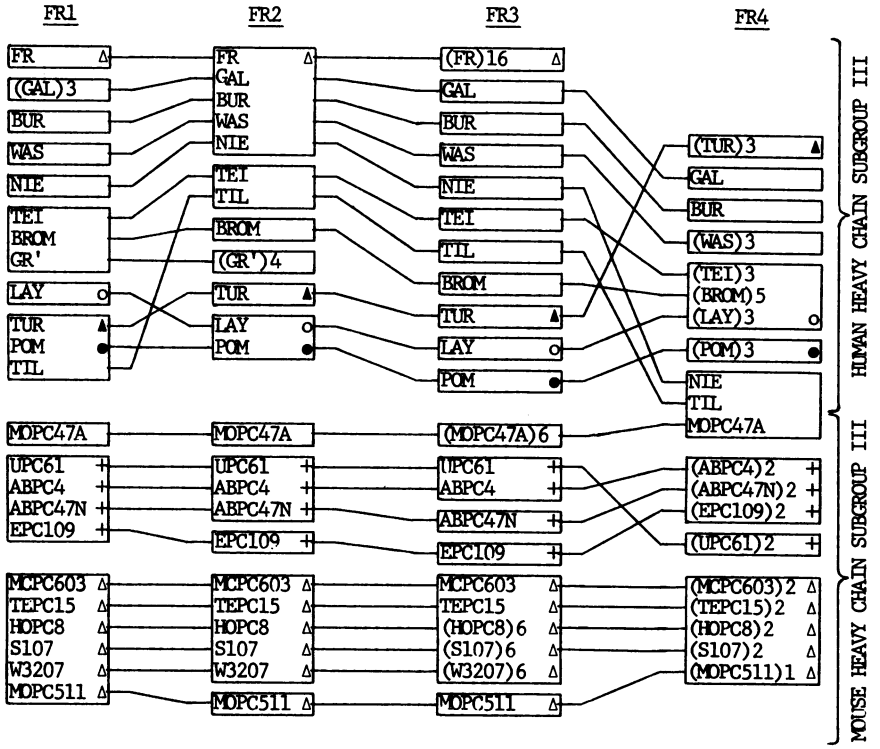

FIG. 4. Assembly of the frameworks of human and of mouse V_HIII chains from minigene sets. ●, Anti-human gamma G1 globulin; ○, anti-human gamma G1 and G3 globulin; ▲, cold agglutinin with anti-blood group I activity; △, anti-phosphocholine; +, anti-$\beta 2 \rightarrow$ 1-fructosan.

If one examines the nucleotide sequences of Tonegawa *et al.* (12) for their V_λ mouse gene in terms of this hypothesis, it is clear that FR1 and FR3 are of V_λII origin, the sequences corresponding to MOPC 315, whereas FR2 has a Val at position 36 (numbering as in ref. 2), the residue found by amino acid sequencing in V_λI, MOPC 104E. The nucleotide sequence for Val^{36} is followed by the nucleotide sequence for Gln-Glu, the amino acids at residues 37 and 38 which are considered to be the V_λII sequence; Gln-Gln is given for V_λI. However, these sequences have not been thoroughly documented with respect to the presence of the amide group. Indeed, the data of Dugan *et al.* (51) give Glx-Glx and those of Goetzl and Metzger (52) give Glu-Glu for the sequence at positions 37 and 38, not Gln-Glu as reported by Tonegawa *et al.* (12). This sequence Gln-Gln in MOPC 104E (53) must also be considered as tentative. This is considered to be a very difficult region to sequence. If the Gln or Glu assignment from amino acid sequencing does not ultimately distinguish between V_λI and V_λII, or if it does not indicate that this segment is V_λII, then FR2 with Val at position 36 becomes a V_λI whereas FR1 and FR3 are clearly V_λII, which would strongly support the hypothesis presented. It thus becomes of substantial importance to determine these sequences unequivocally. Moreover, it is crucial to recognize that nucleic acid sequencing will not replace amino acid sequencing in studying immunoglobulin chains and antibody complementarity until one can be sure which nucleotide sequences, when translated into amino acid sequences, actually are in the protein and which are intervening nucleotides.

The various sets in a given FR may differ from one another by substitutions at from 1 to 11 positions. The high frequency of sets with multiple members, including members of several species, would tend to support the hypothesis that the minigenes are maintained in the germ line rather than that they are generated by somatic mutation which is usually used to explain one or a few amino acid substitutions. As new clones are discovered and sequenced, the association of FR1 with FR2, etc., from different sets, as is a possibility from the first clone studied (12), could also provide a better test of this hypothesis.

If this hypothesis proves to be correct, the location of the CDR in relation to the FR in the genome becomes of great importance and this would have to be studied in sperm DNA. Because recent studies show the apparent linkage of idiotype (which is related closely to the CDR) to allotype to be dependent in the mouse upon the light chain (54), this association becomes essentially a regulatory phenomenon and would not necessarily provide any insight into the location of the CDR or even the FR in relation to the C-region of the genome. Previous findings (55) of an identical CDR1 in a human V_λII and a human V_λV although they differ at 21 positions throughout the V-region and an identical CDR3 (31) in a human V_κI and a human V_κIII despite 30 amino acid differences in the V-region also strongly support the assembly of the V-region by the joining with or insertion of the CDR and the FR pieces (1). In the rabbit there is evidence for linkage of the *a* allotypes, which are in the V-framework, with the C-region.*

Assembly of the V-region framework by selection of different FR pieces would make for a geometric increase in the number of frameworks for the complete V-region. It would also provide a substantial protective mechanism for the antibody-forming system because losses of the minigenes for FR pieces from one or more segments would not prevent the assembly of immunoglobulin V-regions. Species differences such as those seen in individual FR between human and mouse could be the result, during evolutionary time, of loss of FR pieces or of mutations that did not affect the ability of these FR segments to assembly and fold into the proper three-dimensional structure. That four

* Mage, R. G. (1977) *Proceedings of the Third International Congress of Immunology*, Sydney, Australia, July 3–8, *Progress in Immunology*, Vol. 3, pp. 289–297.

sets of human FR4 comprise members of the different V_κ subgroups and that one set of FR4 contains an identical sequence from two species and another set of identical FR2 contains members of three species, indicating substantial conservation of the minigenes for the FR segments in evolution, is crucial to complete understanding of the genetic basis of antibody structure and the generation of diversity.

We have searched for sequences of the minigene sets in the data banks of the C-region sequences and in the Dayhoff file (56) with immunoglobulins removed. The greatest homology found was only 6/11 residues in the FR4 set with two mouse and one human V_HIII. All other sets showed lower degrees of matching. Thus, the proposed minigene sets are unique to V-regions.

The data base of variable region sequences is maintained in the PROPHET computer system (57). This work was sponsored by the National Cancer Institute, National Institute of Allergy and Infectious Diseases, National Institute of Arthritis, Metabolism and Digestive Diseases, and National Institute of General Medical Sciences, Division of Research Resources (Contract no. N01-RR-4-2147), of the National Institutes of Health. T.T.W. is a Career Research Development Awardee (5-K01-AI-70497), National Institutes of Health. The work was aided in part by a grant from the National Institutes of Health, 1-R01-GM21482-01, to T.T.W. and by a grant from the National Science Foundation, BMS-76-81029, to E.A.K.

1. Wu, T. T. & Kabat, E. A. (1970) *J. Exp. Med.* **132,** 211–250.
2. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1976) *Variable Regions of Immunoglobulin Chains* (Bolt, Beranek and Newman, Cambridge, MA).
3. Kabat, E. A. & Wu, T. T. (1971) *Ann. N.Y. Acad. Sci.* **190,** 382–392.
4. Kabat, E. A. (1976) *Structural Concepts in Immunology and Immunochemistry* (Holt, Rinehart and Winston, New York), 2nd Ed.
5. Milstein, C. (1967) *Nature* **216,** 330–332.
6. Nilla, H. & Edman, P. (1967) *Nature* **216,** 262–263.
7. Hood, L., Gray, W. R., Sanders, B. G. & Dreyer, W. J. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32,** 133–145.
8. Barstad, P., Farnsworth, V., Weigert, M., Cohn, M. & Hood, L. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 4096–4100.
9. Capra, J. D. & Kehoe, J. M. (1975) *Adv. Immunol.* **20,** 1–40.
10. Potter, M. (1978) *Adv. Immunol.* **25,** 141–211.
11. Tonegawa, S., Brack, C., Hozumi, N. & Schuller, R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3518–3522.
12. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1485–1489.
13. Gilbert, W. (1978) *Nature* **271,** 501.
14. Tilghman, S. M., Tiemeier, D. C., Polsky, F., Edgell, M. H., Seidman, J. G., Leder, A., Enquist, L. W., Norman, B. & Leder, P. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4406–4410.
15. Tilghman, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P. & Weissman, C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1309–1313.
16. Jeffreys, A. J. & Flavell, R. A. (1977) *Cell* **12,** 1097–1108.
17. Breathnach, R., Mundel, J. H. & Chambon, P. (1977) *Nature* **270,** 314–319.
18. Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1299–1303.
19. Berget, S. M., Moore, C. & Sharp, P. A. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3171–3175.
20. Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. (1977) *Cell* **12,** 1–8.
21. Klessig, D. F. (1977) *Cell* **12,** 9–21.
22. Kitchingman, G. R., Lai, S.-P. & Westphal, H. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4392–4395.
23. Celma, M. L., Dhar, R., Pan, J. & Weissman, S. M. (1977) *Nucleic Acids Res.* **4,** 2549–2559.
24. Aloni, Y., Dhar, R., Laub, O., Horowitz, M. & Khoury, G. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3686–3690.
25. Crawford, L. V., Cole, N. N., Smith, A. E., Paucha, E., Tegtmeyer, P., Rundell, K. & Berg, P. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 117–121.
26. Glover, D. M. & Hogness, D. S. (1977) *Cell* **10,** 167–175.
27. White, R. L. & Hogness, D. S. (1977) *Cell* **10,** 177–192.
28. Wellauer, P. K. & David, I. B. (1977) *Cell.* **10,** 193–212.
29. Pelligrini, M., Manning, J. & Davidson, N. (1977) *Cell* **10,** 213–224.
30. Goodman, H. M., Olson, M. V. & Hall, D. B. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5453–5457.
31. Klapper, D. G. & Capra, J. D. (1976) *Ann. Immunol. (Paris)* **127,** 261–271.
32. Hood, L., Loh, E., Hubert, J., Barstad, P., Eaton, B., Early, P., Fuhrman, J., Johnson, N., Kronenberg, M. & Schilling, J. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **41,** 817–836.
33. Capra, J. D., Klapper, D. G., Tung, A. S. & Nisonoff, A. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **41,** 847–853.
34. Haber, E., Margolies, M. N. & Cannon, L. E. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **41,** 647–659.
35. Braun, D. G. & Huser, H. (1977) "Proceedings of the Third International Congress of Immunology, Sydney, Australia," *Prog. Immunol.* **III,** 255–264.
36. Braun, D. G., Huser, H. & Riesen, W. F. (1976) *Eur. J. Immunol.* **6,** 570–578.
37. Strosberg, A. D., Van Hoegaeden, M. & Schreiber, A. D. (1977) *Ann. Immunol. (Paris)* **128,** 345–349.
38. Freedman, M. H., Guyer, R. B. & Terry, W. D. (1972) *J. Biol. Chem.* **247,** 7051–7061.
39. Capra, J. D. & Hopper, J. E. (1976) *Immunochemistry* **13,** 995–999.
40. Riesen, W. S., Braun, D. G. & Jaton, J. C. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2096–2100.
41. Wang, A. C., Wang, I. Y. & Fudenberg, H. H. (1977) *J. Biol. Chem.* **252,** 7192–7199.
42. Liu, V. Y. S., Low, T. L. K., Infante, A. & Putnam, F. W. (1976) *Science* **193,** 1017–1020.
43. Robinson, E. A. & Appella, E. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 2465–2469.
44. Potter, M., Rudikoff, S., Vrana, M., Rao, D. N. & Mushinski, E. R. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **41,** 661–666.
45. Vrana, M., Rudikoff, S. & Potter, M. (1977) *Biochemistry* **16,** 1170–1174.
46. Potter, M., Rudikoff, S., Padlan, E. & Vrana, M. (1977) in *Antibodies in Human Diagnosis and Therapy* (Raven, New York), pp. 9–28.
47. Kabat, E. A. (1968) *Proc. Natl. Acad. Sci. USA* **59,** 613–618.
48. Capra, J. D. & Kindt, T. J. (1975) *Immunogenetics* **1,** 417–427.
49. Campbell, A. M. (1969) *Episomes* (Harper and Row, New York).
50. Phillips, S. G., Phillips, D. M., Kabat, E. A. & Miller, O. J. (1978) *In Vitro,* in press.
51. Dugan, E. A., Bradshaw, R. A., Simms, E. S. & Eisen, H. N. (1973) *Biochemistry* **12,** 5400–5416.
52. Goetzl, E. J. & Metzger, H. (1970) *Biochemistry* **9,** 3862–3871.
53. Appella, E. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 590–594.
54. Laskin, J. A., Gray, A., Nisonoff, A., Klinman, N. & Gottlieb, P. D. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4600–4604.
55. Wu, T. T., Kabat, E. A. & Bilofsky, H. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 5107–5110.
56. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Springs, MD) [including Supp. I (1973) and Supp. II (1976)].
57. Raub, W. F. (1974) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **33,** 2390–2392.