Exon Shuffling Generates an Immunoglobulin Heavy Chain Gene

Richard Maki; Andre Traunecker; Hitoshi Sakano; William Roeder; Susumu Tonegawa


Stable URL:
http://links.jstor.org/sici?sici=0027-8424%28198004%2977%3A4%3C2138%3AESGAIH%3E2.0.CO%3B2-X

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.
Exon shuffling generates an immunoglobulin heavy chain gene

(gene cloning/gene rearrangement/intron and exon/R-loop mapping/heavy chain switch)

RICHARD MAKI, ANDRÉ TRAUNECKER, HITOSHI SAKANO, WILLIAM ROEDER, AND SUSUMU TONEGAWA

Basel Institute for Immunology, 487 Grenzacherstrasse, Postfach, 4005 Basel, Switzerland

Communicated by Niels Kaj Jerne, December 26, 1979

ABSTRACT From endonuclease EcoRI partial libraries of DNAs from mouse embryo and MOPC 141, a γ2b-producing myeloma, clones were isolated by using a DNA fragment carrying the γ2b constant (C) region gene as a hybridization probe. One clone from MOPC 141 contained a heavy chain variable (V) gene and the Cγ2b gene, as demonstrated by R-loop mapping. The V gene and C gene in this clone were separated by a 3.9-kilobase intron. The characterization of this clone as well as the embryonic clones suggest that at least two recombination events occurred to create the γ2b gene in MOPC 141. One of the events is DNA joining to the V-J joining previously demonstrated in the light chain genes, which brings the major part of the V gene next to a short coding sequence (J). The other event we refer to as “Cμ-Cγ2b switch recombination” because a portion of the intron between the V gene and C gene of the rearranged γ gene is derived from the 5′ flanking sequence of the embryonic Cμ gene. A model suggesting how the phenomenon of switch seen in lymphocytes may occur is presented.

The notion that DNA segments in mouse embryo are rearranged during the ontogeny of a bone marrow-derived (B) lymphocyte to create a Λ or κ immunoglobulin light chain gene has been demonstrated (1, 2). Essentially, this process involves bringing together DNA segment coding for one of multiple variable (V) regions with a DNA segment coding for the corresponding constant (C) region. The detailed studies of the sequence organization before and after recombination have revealed that the V region is actually encoded in two separate DNA segments (3, 4). For example, in the λ system the major part of the V region (residues 1–97) is encoded in the V DNA segment, whereas the rest of the V region (residues 98–109) is encoded in a short DNA segment referred to as J DNA. A recombination event between the 5′ end, relative to the direction of transcription, of the V DNA segment and the 5′ end of the J DNA segment joins these DNA segments in the B lymphocyte. The J DNA segments have been mapped to the 5′ side of both Cλ and Cκ genes. Whereas in the λ system a single J DNA segment has been found 1.2 kilobases (kb) away from CΛ (3, 5), in the κ system there is a cluster of at least five J DNA segments in the range of 2.5 to 4.4 kb away from the Cκ gene (6). The possibility that the J DNA segment, in addition to providing a recombination point for any V DNA, may also contribute to the antibody diversity has been suggested and may explain the need for a cluster of J DNA segments (6, 7).

Because the heavy chain polypeptide is also composed of a V region and a C region, one might assume that what has been established at the DNA level for light chains will be true for heavy chains. In principle, this is probably so, although the expression of heavy chains has some features that make it considerably more complex than that of light chains. The nature of this complexity is revealed by the existence of at least eight different C regions. Each of these C regions appears to share the same set of V regions; i.e., a given V gene can be expressed with more than one heavy chain class. In addition, it appears that within a single lymphocyte there is the capacity to switch the class of C gene expressed while maintaining expression of the same V region (8–10).

In this report we present several features of the organization of heavy chain genes in both mouse embryo and a γ2b-synthesizing myeloma, MOPC 141. These include: (i) the distance separating two heavy chain C genes, (ii) evidence that at least two rearrangements at the DNA level are involved in the generation of a complete immunoglobulin gene of the γ2b subclass, (iii) the intron–exon structure of the embryonic Cμ gene, and (iv) the existence of a J coding DNA segment for the heavy chain genes.

MATERIALS AND METHODS

Bacteria and Phages. Phage Charon 4A and Escherichia coli DF50 (Su II, Su III′) were obtained from F. Blattner of the University of Wisconsin, Madison, WI (11). The AgtWES-AB was obtained from P. Leder at the National Institutes of Health (12). E. coli 802 (Su, m3, Su III′) was originally from K. and N. Murray of the University of Edinburgh. Lysozyme used for preparation of packaging mixtures, BHB 2698 [N205 recA− (λ imm434 b2 red3 E4m4 Sam7)/λ] and BHB 2690 [N205 recA− (λ imm434 c1ts b2 red3 Dami15 Sam7)/λ] were obtained from B. Hohn of the Friedrich Miescher Institute (13).

Preparation of Mouse Embryo and MOPC 141 Libraries. High molecular weight DNA from either mouse embryos or MOPC 141 tumors was prepared as described (14). The DNA was partially digested with restriction endonuclease EcoRI and fractionated on a 10–40% sucrose gradient. DNA fragments in the size range of 10–20 kb were ligated to Charon 4A arms and packaged in vitro; the efficiency of packaging was 5 × 106 recombinant phage per μg of mouse DNA (13). Each library was amplified as a plate lysate before screening.

Isolation of Clones. Nick-translated hybridization probes were prepared as described (3). Approximately 600,000 plaques from each library were screened as discussed elsewhere (15).

Electron Microscopy. Procedures used for formation of single-stranded DNA-mRNA hybrids, formation of DNA heteroduplexes, two-step R-loop technique, and formation of R-hybrids have been described (16). In the two-step R-loop technique the R-loop mixture is completely denatured by heating and then incubated at 57°C for 4–6 hr. The R-hybrid technique is essentially the same as the two-step R-loop technique except that two different cloned DNAs are mixed with mRNA.

Other Procedures. Purification of γ chain mRNAs of γ1 (MOPC 21), γ2b (MOPC 173), γ2b (MOPC 141), and γ3 (J 606),

Abbreviations: V and C, variable and constant regions of immunoglobulins; kb, kilobase(s); bp, base pair(s).

2138
RESULTS

Cγ1 and Cγ2b genes are 21 kb apart in the mouse embryo

We previously cloned from EcoRI-digested mouse embryo DNA a 6.8-kb fragment carrying a Cγ1 gene (18). A 3.2-kb Xho I/Kpn I fragment (Fig. 1) containing the Cγ1 gene was isolated from this fragment and used as a hybridization probe to screen the embryo EcoRI partial library. Isolated clones were analyzed by EcoRI digestion followed by agarose gel electrophoresis. We also examined R-loops formed between each of these clones and a series of mRNAs coding for different heavy chain classes and subclasses. These studies permitted us to classify all clones into two groups. One group contained the Cγ1 gene on a 6.8-kb fragment, 3.2 kb from the 5' EcoRI site and 1.9 kb from the 3' EcoRI site; the other group contained the Cγ2b gene on a 6.6-kb fragment, 1.4 kb from the 5' EcoRI site and 3.8 kb from the 3' EcoRI site. Clones of the latter class were isolated with the Cγ1 gene probe presumably by virtue of the cross-hybridizability between the Cγ1 gene probe and the Cγ2b gene (Fig. 1).

The order of EcoRI fragments in each clone was determined by R-hybrid analysis. For example, an R-hybrid mixture containing MEP 15 and MEP 1 plus γ1 mRNA revealed that both clones were completely homologous except for a 5.1-kb segment. The position of the R-loops for the Cγ1 gene allowed us to conclude that the 5.1-kb segment is 8.2 kb away from the 3' end of the Cγ1 gene in the 3' direction. Thus, the order of the EcoRI fragments in MEP 15 is a 6.8-kb fragment that contains the Cγ1 gene, followed by a 6.3-kb fragment to the 3' side of the gene, and next to this fragment a 5.1-kb fragment.

Digestion of one of the Cγ1-containing clones, MEP 15, and one of the Cγ2b-containing clones, MEP 5, with EcoRI demonstrated that both clones contained a 5.1-kb EcoRI fragment (Fig. 1). Heteroduplex analysis of the two clones revealed that the two 5.1-kb fragments were indistinguishable. On the basis of this result, it was concluded that the Cγ1 gene was to the 5' side of the Cγ2b gene and the two genes were separated by 21.1 kb. In several of the Cγ2b-containing clones, mapping studies revealed a small deletion ranging in size from 1 to 2 kb as shown for clones MEP 5 and MEP 8 in Fig. 1. These clones constituted a minority of the Cγ2b-containing clones and probably resulted during the growth of the phage. No further attempt was made to characterize this deletion.

γ2b gene is rearranged in MOPC 141

The 6.6-kb EcoRI fragment containing the Cγ2b gene was used as a hybridization probe to screen the EcoRI partial library of MOPC 141 myeloma, a γ2b producer. Of the six clones isolated, two were identical and contained two EcoRI fragments, an 8.2-kb fragment containing the MOPC 141 V gene and a 7.8-kb
fragment containing the CY<sub>2b</sub> gene. A hybrid formed between one of these two clones, 141-p21 and MOPC 141 γ<sub>2b</sub> mRNA, is shown in Fig. 2. The two genes are separated by a 3.9-kb intron. The CY<sub>2b</sub> gene is composed of four exons; the three largest—E<sub>1</sub> (256 bp), E<sub>2</sub> (327 bp), and E<sub>3</sub> (392 bp)—are interpreted to correspond to the C region domains of a γ<sub>2b</sub> chain, whereas the short E<sub>4</sub> (56 bp) very likely codes for the hinge region. This structure is similar to that observed previously for the CY<sub>1</sub> gene (18).

The R-hybrid formed between the myeloma clone, 141-p21, and the embryonic CY<sub>2b</sub> gene clone, MEP 3, is shown in Fig. 3a. The two cloned DNA fragments are extensively homologous near their 3′ ends, but, unlike the light chain genes, the homology does not extend to the V gene, but only 1.3 kb from the 5′ end of the C gene into the intron between the V gene and C gene (Fig. 3c). This result suggests that, although the myeloma γ<sub>2b</sub> gene is rearranged, the rearrangement is different from the way the light chain genes are rearranged.

Embryonic DNA fragment containing the remainder of the 3.9-kb intron

The V–C intron in 141-p21 was found to be 3.9 kb long, of which 1.3 kb could be accounted for by the homology with MEP 3. To determine the origin of the 2.6-kb intron segment unaccounted for by the embryonic CY<sub>2b</sub> gene clone, MEP 3, we isolated from clone 141-p21 a 1.2-kb Sac I/EcoRI fragment. The location of this fragment in 141-p21 is shown in Fig. 2. Five clones were isolated by using this fragment as a hybridization probe to screen the mouse embryo library. One of these clones, MEP 203, contained three EcoRI fragments. Heteroduplex analysis of the clones allowed us to order the three EcoRI fragments as follows: 1.5 kb, 6.4 kb, and 10.5 kb. The R-hybrid of MEP 203 and the myeloma clone 141-p21 is shown in Fig. 3b. The γ<sub>2b</sub> mRNA was added to determine the position of the homologous region relative to the CY<sub>2b</sub> gene. On the myeloma clone there is a homologous segment between 1.3 kb and 3.9 kb to the 5′ side of the CY<sub>2b</sub> gene. Because the V and the CY<sub>2b</sub> gene are separated by a 3.9-kb intron (Fig. 2), the 5′ end of the homology corresponds to the 3′ end of the V gene on 141-p21 (Fig. 3c). Thus, part of the intron between the V gene and C gene in the myeloma clone 141-p21 is due to a segment of DNA on MEP 203 and the rest of the intron can be found on clone MEP 3 (Fig. 3c).

Embryonic clone MEP 203 contains a C<sub>μ</sub> gene

R-loop analysis of EcoRI-digested clone MEP 203 demonstrated that this clone contained a C<sub>μ</sub> gene on the 10.5-kb EcoRI fragment. Partially purified μ heavy chain mRNA from MOPC 104E gave four R-loops, each separated by short single-stranded DNA segment (Fig. 4). The approximate lengths of the exons are: E<sub>3</sub>, 343 bp; E<sub>2</sub>, 350 bp; E<sub>3</sub>, 359 bp; and E<sub>4</sub>, 506 bp. The multiple exon structure is analogous to the structures previously observed for other heavy chain C genes and strongly suggests that the four C<sub>μ</sub> domains are encoded separately. We also used γ<sub>1</sub>, γ<sub>2a</sub>, γ<sub>2b</sub>, γ<sub>4</sub>, and α heavy chain mRNAs for the analysis and found no R-loops on clone MEP 203.
The 6.4-kb embryonic fragment contains a MOPC 141 J DNA segment

The R-hybrid structure shown in Fig. 3b suggests that one of the two branch points between the myeloma clone 141-p21 and the embryo clone MEP 203 corresponds to the 3′ end of the V gene. If there is a V-J joining event in heavy chain genes similar to those seen for the light chain genes, then the 6.4-kb embryonic fragment should contain a J DNA segment at a position corresponding to the branch point. The J DNA segment is too short to be visualized as an open R-loop, but its presence can be inferred by a duplex DNA loop formed by the intron separating the embryonic J and C DNA segments.

In order to determine whether a J DNA segment coding for the MOPC 141 γ2b chain is present on the 6.4-kb fragment, this DNA fragment was ligated to the Cγ2b gene-positive 6.6-kb EcoRI fragment of clone MEP 3 (Fig. 1) and recloned in the λ WES vector. Head-to-tail ligation of the two EcoRI fragments in one of the clones isolated (clone ME 46) was verified by measurement of various parts of the heteroduplex structure obtained with this clone and the myeloma clone 141-p21 (data not shown). The R-loop structure formed by clone ME 46 and partially purified MOPC 141 γ2b mRNA is shown in Fig. 5. The 2.7-kb single-stranded DNA loop (labeled in Fig. 5 as IA) indicates the presence of a J DNA segment 2.7 kb away from the 5′ side of the Cγ2b gene. Because the Cγ2b gene is 1.4 kb away from the 5′ EcoRI end of the 6.6-kb fragment (Fig. 5), the above result places the J DNA segment on the 6.4-kb fragment and 1.3 kb away from its 3′ EcoRI end. This position corresponds to one of the fork points of the heteroduplex formed between clone MEP 203 and clone 141-p21 (see Fig. 3c). Recent DNA sequence analysis of this fork point in clone 141-p21 and MEP 203 has revealed that a J DNA segment similar to that identified for light chains does exist for the heavy chains (unpublished observations). These results suggest that the myeloma heavy chain V gene is generated by V-J joining.

DISCUSSION

The results presented in this report suggest that at least two recombination events are necessary for generation of the complete γ2b gene from embryonic DNA sequences. One of them occurs at or near the 3′ end of the V-coding sequence and a short (∼50 bp) sequence that seems to code for the V-C junction region. The latter sequence is analogous to the J DNA segment of light chain genes (3). We designate this sequence as a heavy chain J DNA. In contrast to the light chain genes, in which J DNA segments are located in the 5′ flanking regions of their respective C genes, in the heavy chain gene system the J DNA used for the MOPC 141 γ2b gene does not lie in the 5′ flanking region of the Cγ2b embryo gene but is located in the 5′ flanking region of the Cc embryo gene. Thus, there may only be one set of J DNA segments for the entire set of heavy chain C genes.
The second recombination occurs between a pair of sites, one located between the J DNA segment and the Cμ gene and the other in the 5' flanking sequence of the Cγ2b gene. This event we refer to as "Cμ-Cγ2b switch recombination." The temporal order of the two recombinations during B cell differentiation cannot be determined from the present studies. However, the attractive possibility, in light of the heavy chain "switch" phenomenon (8–10), is that V-J joining occurs first, followed by Cμ-Cγ2b switch recombination (Fig. 6). According to this hypothesis the IgM to IgC2b switch can be explained in the following manner. During the differentiation of stem cells to pre-B cells, one of the multiple V DNA segments is joined to one of the J DNA segments located to the 5' side of the Cμ gene. In analogy to the light chain genes, it is probable that this recombination would generate a complete and active μ gene. The μ-only pre-B cell (19, 20) and the resting small B lymphocyte bearing IgM on its cell surface (21) are considered to contain μ genes of such sequence organization. Upon encountering an antigen and often with the help of T cells, the small B lymphocytes undergo further differentiation and become blast cells and eventually plasma cells. During this process the second rearrangement, Cμ-Cγ2b switch, occurs in some lymphocytes, replacing the Cμ coding exons of the active μ gene with the Cγ2b exons. Subsequently, these cells synthesize a γ2b chain having the same V region as the μ chain synthesized by their progenitors.

Many eukaryotic genes are composed of informational DNA (exons) interspersed with silent sequences (introns). In the immunoglobulin genes a polypeptide encoded by a single exon seems to form a structurally and functionally independent unit. Thus, the signal peptide and the major portion of the V region are encoded in separate exons (5). Similarly, each of the three heavy chain C region domains, as well as the hinge region of γ1 and γ2b chains, is encoded separately (18, 22, 23). The B-loop analysis of the Cμ gene presented here suggests that this heavy chain gene also follows this pattern. Correlation between a signal peptide and an exon has been shown in at least one other case (24). We (5, 18) as well as others (25, 26) have previously speculated that one function of an intron is to serve as a site of recombination between exons. In this way, two or more DNA segments scattered throughout the genome and coding for a polypeptide of some functional use can be assembled into a single transcription unit. Consequently, a gene coding for a protein having a new combination of functions is created. While it is yet to be shown whether any genes were generated by shuffling of exons in evolution, the structure of the γμ gene from MOPC 141 suggests that such a gene creation process indeed takes place in differentiation.

Note Added in Proof. Recently two small exons have been observed approximately 2 kb from the end of the Cμ gene. These exons may be used for the membrane form of μ.

We thank Dr. M. Potter for providing us with various myeloma cell lines. We also thank Ms. L. Gibson, Ms. P. Riegert, and Mr. G. Dastoorinukko for technical assistance and Ms. S. Buser for the secretarial work.