片岡徹

主論文

Shine H0926 Honjo

<af 509>

'23, 000-000, February 1981" ' 1981"

' Repetitive Sequences in Class-Switch Recombination Regions of Immunoglobulin Heavy Chain Genes"

<af 514>

' Tohru Kataoka, Takashi Miyata* and Tasuku Honjo"

do not break authors' name

' Department of Genetics
Osaka University Medical School
Osaka 530, Japan and
* Department of Biology
Faculty of Science, Kyushu University
Fukuoka 812, Japan"

<af 514>

~~Running Title:~~ ' Immunoglobulin Gene Class Switch" '<rn—>"

Summary

Immunoglobulin class switch involves a unique recombination event that takes place at the region 5′ to each heavy chain constant region gene during B lymphocyte differentiation. Such regions that are responsible for the class-switch recombination are defined as S regions (Kataoka et al., Proc. Natl. Acad. Sci. USA 77, 919, 1980). We have cloned a rearranged γ2b gene from a mouse myeloma (MPC11) and compared its structure with the germ line counterparts. The rearranged γ2b gene contained the 5′ flanking region of the γ3 gene ($S_{\gamma 3}$ region) which are linked to the 5′ flanking region of the γ2b gene ($S_{\gamma 2b}$ region). We have determined nucleotide sequences surrounding the recombination site of the rearranged and germ line γ2b genes, which include the $S_{\gamma 2b}$ and $S_{\gamma 3}$ regions. Both $S_{\gamma 2b}$ and $S_{\gamma 3}$ regions comprise tandem repetition of conserved units of 49 base-pairs. Similar 49 base-pair repeating units are also found in the previously determined sequence of the $S_{\gamma 1}$ region in which class-switch recombination took place in MC101 myeloma. The nucleotide sequences of the $S_{\gamma 1}$, $S_{\gamma 2b}$ and $S_{\gamma 3}$ repeating units share significant homology with each other. The S_{μ} region, partial nucleotide sequence of which was previously determined, contains abundant short sequences such as AGCT, TGGG and AGCTGGGG which are shared in common by repeating sequences in S_{γ} regions. These results suggest that the recombination responsible for class switch from μ to γ or from a γ to another γ, may be facilitated directly or indirectly by homology of repeating sequences in S regions."

Introduction

An immunoglobulin molecule contains one of two types of light (L) chain (κ or λ) and one of eight types of heavy (H) chain (γ, δ, α, γ1, γ2a, γ2b, γ3 and ε) in mouse, the heavy chain defining the immunoglobulin class (IgM, IgD, IgA, IgG1, IgG2a, IgG2b, IgG3 and IgE). Both L and H chains comprise a variable (V) region and a constant (C) region. During differentiation of a given B lymphocyte, a single V gene is first expressed as a part of the μ chain; at a later stage the C region of the expressed H chain switches a C part from μ to γ or to α without alteration of the V region sequence (Nossal et al., 1971; Fu et al., 1975; Gearhart et al., 1975; Goding and Layton, 1976; Sledge et al., 1976; Wang et al., 1977; Abney et al., 1978). Such a phenomenon is called H chain class switch.

We have recently proposed a molecular mechanism to explain H chain class switch based on the comparative studies on the rearranged and germ line H chain genes as shown in Figure 1 (Kataoka et al., 1980; Takahashi et al., 1980). A complete H chain gene is supposed to be formed by at least two types of the recombinational event. The first type of recombination takes place between a given V_H, a J_H and possibly a D gene segment (Early et al., 1980; Honjo et al., 1980; Sakano et al., 1980), completing a V region sequence. After such recombination, referred to as V=J recombination, the V region sequence is expressed as a part of the μ chain. The second type of recombination takes place between S_μ and S_γ (or S_α) regions. The S region was defined as the functional region responsible for the class switch and assumed to be located in the 5′ flanking region of each C_H gene. The S_μ region is located between J_H and the μ gene and the S_γ (or S_α) is present at the 5′ side of each γ gene (or α gene). The second type of recombination, called S=S recombination, can replace the C part of the H chain without affecting the V region sequence. Both types of

recombination result in the deletion of the inbetween DNA segment from the chromosome (Honjo and Kataoka, 1978; Sakano et al., 1979; Coleclough et al., 1980; Cory and Adams, 1980; Cory et al., 1980; Rabbitts et al., 1980; Yaoita and Honjo, 1980a, 1980b). Evidence that a complete H chain gene is formed by at least two recombinational events was also reported by other investigators (Davis et al., 1980a; Maki et al., 1980).

To understand molecular mechanism for the S=S recombination we have undertaken structural analysis of the S_{γ} regions. The nucleotide sequence determination of the $S_{\gamma 1}$, $S_{\gamma 2b}$ and $S_{\gamma 3}$ regions revealed a unique structural feature; the tandem repetition of the 49 bp unit is present in the region flanking the 5' end of these γ genes.

<af 517>

Results

Cloning a Rearranged γ2b Gene from MPC11 Myeloma

High molecular weight DNA extracted from MPC11 myeloma cell nuclei was cleaved partially with Eco RI, and the digests were fractionated on a preparative agarose gel. When an aliquot of each fraction was again electrophoresed in an agarose gel, transferred to a nitrocellulose filter and hybridized with the ^{32}P-labeled Eco RI fragment (6.6 kb) of the γ2b gene cloned from newborn mice (Igγ2b-22) (Kataoka et al., 1979), two dark bands of 13.3 kb and 10.5 kb and one faint band of 23 kb were observed in addition to the 6.6 kb band (data not shown). The 23 kb band is due to the γ2a gene which cross-hybridizes with the γ2b gene as shown previously (Rabbitts et al., 1980; Yaoita and Honjo, 1980a).

The DNA fractions containing the 13.3 kb γ2b gene fragment were pooled and used as inserts for cloning with λgtWES. Two independent phages, which hybridized with ^{32}P-labeled cloned γ2b cDNA (Kataoka et al., 1979), were obtained by screening about 1.5 x 10^5 plaques. Two phages are different and referred to as

λgtWES$_{\lambda}^{\cup}$Igγ2b=2203 and λgtWES$_{\lambda}^{\cup}$Igγ2b=2206, the inserts of which are called ~~as~~ Igγ2b=2203 and Igγ2b=2206, respectively. The cloned phage DNAs were cut with various restriction enzymes and their cleavage sites were mapped by the conventional procedure. The results are summarized in Figure 2. Each of the cloned DNAs consists of the 6.6 and 6.8 kb Eco RI fragments which were recloned into pBR322 (Bolivar et al., 1977) to construct the detailed restriction maps. The plasmids containing the 5′ Eco RI fragments of Igγ2b=2203 and Igγ2b=2206 are called pIgγ2b=23 and pIgγ2b=26, respectively. The 3′ Eco RI fragments of Igγ2b=2203 and Igγ2b=2206 are identical to Igγ2b=22, an embryonic form of the γ2b gene (Kataoka et al., 1979; Yamawaki-Kataoka et al., 1980). Igγ2b=2203 and Igγ2b=2206 seem to differ from each other in their 5′ regions of about 5.2 kb in length. R-loop analysis and Southern blot hybridization using MPC11 (eleven) mRNA failed to find the V gene sequence in either of these clones.

We have also cloned the 10.5 kb fragment hybridizing to the cloned γ2b gene from a Charon 4A phage library containing partial Eco RI digests of MPC11 (eleven) DNA. This clone, referred to as Igγ2b=2204, contains two Eco RI fragments; the 6.6 kb γ2b gene fragment (Igγ2b=22) and a 3.6 kb fragment flanking the 3′ end of Igγ2b=22. Clones containing the same fragments as Igγ2b=2204 were also obtained from mouse embryonic DNA library (Honjo et al., 1980).

<af 771>

Igγ2b-2203 is a Rearranged γ2b Gene and Igγ2b=2206 is the Embryonic γ2b Gene

To determine which of the clones Igγ2b=2203 and Igγ2b=2206 corresponds to the rearranged γ2b gene in MPC11 (eleven) myeloma, we analyzed γ2b gene fragments in DNAs of various myelomas and newborn mice by Southern blot hybridization. We used, as a probe, a 1.3 kb Bam HI fragment of Igγ2b=22 (fragment A in Figure 2) that contains the 5′ flanking region and a portion of the CH1 (one) domain sequence of the γ2b gene. DNAs of newborn mice, various myeloma cells, λgtWES$_{\lambda}^{\cup}$Igγ2b=2206 and λgtWES$_{\lambda}^{\cup}$Igγ2b=2203 were cut with several restriction enzymes ~~which~~ that can

distinguish between Igγ2b=2206 and ~~from~~ Igγ2b=2203. DNA digests were electrophoresed in 1% agarose gels, blotted to nitrocellulose filters and hybridized with the nick-translated fragment A.

Newborn mouse DNA produced the γ2b gene fragments of 6.6, 6.5 and 4.2 kb upon digestion with Sac I, Hinc II and Kpn I, respectively, as shown Figure 3 (lane a). These fragments of newborn mouse DNA are identical to those produced by digestion of Igγ2b=2206 DNA with the homologous enzymes (lane j). Digestion of MPC11 (eleven) DNA with Sac I, Hinc II and Kpn I yielded 3.8, 8.1 and 4.2 kb fragments, respectively, which hybridized with fragment A (lane g). These fragments of MPC11 (eleven) DNA are identical to those produced by digestion of Igγ2b=2203 DNA with the homologous enzymes (lane k). The results indicate that MPC11 (eleven) DNA contains a single form of the γ2b gene that is identical to Igγ2b=2203. We will discuss below why we cloned two types of γ2b genes from MPC11 (eleven) DNA. It is likely that Igγ2b=2203 is a 3′ portion of an expressed γ2b gene in MPC11 (eleven) myeloma although there is no MPC11 (eleven) V region sequence found in this clone. We have recently demonstrated by cloning overlapping DNA segments that the MPC11 (eleven) V gene ~~is located about 16 kb~~ do exist 5′ to the 5′ end of Igγ2b=23 (T. Kataoka, unpublished data). Newborn mouse DNA contains the embryonic form of the γ2b gene that is equivalent to Igγ2b=2206. Taken together, the results indicate that the rearranged gene Igγ2b=2203 was formed by the replacement of the 5′ portion (5.2 kb) of the germ line gene Igγ2b=2206 with a new DNA segment.

When newborn mouse DNA and various myeloma DNAs were digested with Sac I, Hinc II and Kpn I, we found faint bands of 3.9, 5.3 and 7.6 kb, respectively, which are common to all the DNAs except for α chain-producing myeloma DNAs. These bands seem to be attributable to cross-hybridization of the probe to γ2a gene fragments (Y. Yamawaki-Kataoka and T. Honjo, unpublished data).

The above conclusion is fortified by direct cloning of a 6.8 kb Eco RI

fragment of newborn mouse DNA which hybridized with Igγ2b=23 sharing the 3′ 1.5 kb region in common with Igγ2b=26. All ~~of~~ eight clones obtained from the Eco RI digests of newborn mouse DNA were Igγ2b=26 (N. Takahashi and T. Honjo, unpublished data), indicating that Igγ2b=26 is a germ line gene. Furthermore we have cloned the identical DNA fragment to Igγ2b=2206 from a partial Hae III digestion library of embryonic mouse DNA (Honjo et al., 1980; Shimizu et al., 1980 ~~manuscript submitted~~ # for # publication).

In view of the fact that MPC11 (eleven) DNA contains the rearranged γ2b gene of Igγ2b=2203 type almost exclusively (Figure 3, lane g), it is puzzling for us to have cloned not only Igγ2b=2203 but also Igγ2b=2206, the embryonic form of the γ2b gene from MPC11 (eleven) DNA. A most reasonable explanation is that Igγ2b=2206 is derived from DNA of host cells which contaminate MPC11 (eleven) myeloma tumor. Such contamination is obvious by the presence of a faint band equivalent to Igγ2b=2206 in MPC11 (eleven) DNA, and was also shown in other myeloma tumors (Yaoita and Honjo, 1980a). Although it is rare to clone a minor gene fragment in the first two clones, it is not impossible.

DNAs of μ and γ3 chain-producing myeloma contained the γ2b gene fragment identical to the newborn mouse DNA (Figure 3, lanes b and c). DNAs of γ2b chain-producing myeloma contained rearranged γ2b gene fragments (lanes f and g). The γ2b gene fragments were deleted from both chromosomes of γ2a and α (Grk. l.c. alpha) chain-producing myelomas (lanes h and i). These results are in agreement with the deletion model and the order of the C_H genes as summarized in Figure 1 (one) (Honjo and Kataoka, 1978; Coleclough et al., 1980; Cory and Adams, 1980; Cory et al., 1980; Kataoka et al., 1980; Rabbitts et al., 1980; Yaoita and Honjo, 1980a, # 1980b). Unexpectedly, DNAs of γ1 (one) chain-producing myelomas contained the γ2b gene fragment of a reduced size (lanes d and e). Both Sac I and Hinc II fragments in these DNAs were shortened by the same length (about 2.2 kb) ~~as~~ compared with those in newborn mouse DNA.

The Rearranged γ2b Gene Contains a DNA Segment Flanking the Germline γ3 Gene

To search for the origin of a new DNA segment introduced into the 5′ portion of Igγ2b=2203, we analyzed Eco RI digests of newborn mouse DNA by Southern blot hybridization. When the filter was hybridized with ^{32}P=labeled pIgγ2b=23 as a probe, we found 18 kb and 6.8 kb bands in Eco RI digest of newborn mouse DNA as shown in Figure 4A. The 6.8 kb band is equivalent to Igγ2b=26. MPC11 DNA contained a 6.8 kb fragment ~~which~~ that is equivalent to Igγ2b=23 but did not contain the 18 kb fragment. The results suggest that a segment of the 18 kb fragment and the embryonic γ2b gene fragment (Igγ2b=2206) recombined to form the rearranged γ2b gene in MPC11 DNA (Igγ2b=2203).

The 18 kb fragment in Eco RI digests of newborn DNA was cloned using Charon 4A as an EK2 vector. The restriction enzyme cleavage map of the 5′ region (5.2 kb) of this clone is indistinguishable from that of Igγ2b=23 as shown in Figure 5. The restriction map of the other part of the cloned 18 kb fragment is almost identical to the 3′ part of the γ3 gene fragment cloned from DNA of the γ3 chain=producing myeloma J606 (Y. Yamawaki=Kataoka and T. Honjo, unpublished data), suggesting that the cloned 18 kb fragment is the γ3 gene fragment.

To test this possibility, the cloned 18 kb fragment was digested with Eco RI, Hind III + Eco RI and Xba I + Eco RI and hybridized with the nick-translated cloned γ3 cDNA according to Southern's (1975) procedure. As shown in Figure 4B, the ^{32}P=labeled cloned γ3 cDNA hybridized with a specific segment of the 18 kb fragment cloned from newborn mouse DNA, which was thereafter referred to as Igγ3=30. These results unequivocally demonstrate that the rearranged γ2b gene (Igγ2b=2203) contains, at the 5′ side of the γ2b gene, the DNA segment derived from the 5′ flanking region of the γ3 gene. Igγ2b=2203 appears to be formed by the class=switch recombination between the $S_{\gamma 3}$ and $S_{\gamma 2b}$ regions, indicating that the H chain class switch can take place from a γ to another γ.

<af 771>

Nucleotide Sequence surrounding Recombination Site of the γ2b Gene"

To understand the molecular mechanism for the class-switch recombination, we have determined nucleotide sequences of the regions (about 2kb) surrounding the recombination site in Igγ2b-23 and Igγ2b-26. Figure 2 illustrates the detailed restriction maps of those regions together with the strategies used for sequencing. All the restriction cleavage sites of Igγ2b-26 and Igγ2b-23 were identical in the region downstream from a recombination site shown in Figure 2, except that a Hae III site is present at about 200 bp 3′ to the recombination site of Igγ2b-23 but not at the corresponding position of Igγ2b-26.

Comparison of the nucleotide sequences of Igγ2b-26 and Igγ2b-23 unambiguously determined the recombination site at 35 bp 3′ side of the Sac I site of Igγ2b-23, which is located 5.2 kb from the 5′ end of Igγ2b-2203 as shown in Figure 6. The nucleotide sequences of Igγ2b-26 and Igγ2b-23 matched each other almost completely in the region 3′ to the recombination site except for a transition of A to G at base 217 of Igγ2b-23, which explains the unique Hae III cleavage site of Igγ2b-23. The discrepancy may be ascribed to a point mutation in MPC11 (eleven) myeloma.

It is important to point out that no sequences that could code for the amino acids of J segments of H chain V regions (Rao et al., 1979) were found anywhere surrounding the recombination site of Igγ2b-23 and Igγ2b-26, indicating that the nature of the recombination in Igγ2b-23 was quite different from that of the V-J recombination in L and H chain genes (Sakano et al., 1979 ②; Max et al., 1979 ①; Early et al., 1980; Honjo et al., 1980; Sakano et al., 1980) as shown previously (Kataoka et al., 1980; Takahashi et al., 1980).

<af 771>

Tandem Repetition of 49 bp Unit in the Sγ2b Region"

The whole region sequenced in Igγ2b-26 consists of a tandem array of 49 bp repeat units as shown in Figure 7. The prevalent nucleotide sequence of the 49

Waverly - underline as typed

bp repeat unit was represented by GGGACCAG$^{T}_{A}$CCTAGCAGCTPuTGGGGGAGCTGGGGA$^{A}_{T}$GGTPuGGA PuTPuTGA. Each unit shares at least 80% homology with the prevalent sequence and usually includes two Alu I cleavage sites (AGCT) and one Ava II cleavage site (GGACC), the latter being periodically altered to a Kpn I cleavage site (GGTACC). There are unanimously conserved positions within 49 bp unit; G at positions 17 and 22; A at positions 7, 13, 16, 27 and 43; T at positions 39 and 47; and C at position 18.

Furthermore, based on the sequence homology, one can divide these repeating units into six tandem groups, each consisting of five or six 49 bp repeating units. The groups were numbered from 5′ to 3′ and aligned to maximize homology (Figure 7). The first repeating group starts probably at base minus 819 and the sixth group ends at base 798. The nucleotide sequences of the six repeating groups were similar to each other, amongst which the two 294 bp groups (groups one and 4 in Figure 7) share 94% of the nucleotide sequence in common.

The range of the repetitive sequences in Igγ2b=26 was determined by partial restriction cleavage with Alu I and Ava II, and partial nucleotide sequence determination of the neighboring DNA fragments. A 1.7 kb ~~long~~ Hinc II=Xba I fragment of Igγ2b=26 (fragment B in Figure 2) was labeled with ^{32}P only at the Xba I cleavage site, ~~which lied~~ at the 5′ side of the region sequenced (Figure 6). The ^{32}P=labeled fragment B was cleaved partially with Ava II or Alu I for various time intervals, electrophoresed in 5% polyacrylamide gels and autoradiographed. The partial digestion pattern shown in Figure 8 indicates that the Ava II cleavage sites are located 105, 152, 187, 305, 360, 410, 560, 615, 680 and 780 bp away from the Xba I site and that the Alu I cleavage sites are present at 20, 47, 81, 92, 130, 195, 230, 320 and 395 bp apart from the Xba I site. The combination of the Ava II cleavage sites (GGACC) shown above and the known Kpn I cleavage sites (GGTACC) at approximately 10, 260 and 510 bp 5′ side from the

labeled Xba I site (Figure 2) suggests a regular repetition of these restriction sites ~~with~~ at about 50 or 100 bp intervals. There seems to be multiple Ava II cleavage sites at around 187 bp away from the Xba I site because this band does not disappear after long digestion (Figure 8). The Alu I cleavage sites (81, 92, 130, 195, 230, 320 and 395 bases from the Xba I site) appear ~~with similar~~ at intervals similar to those present between positions -105 and -415 of Figure 6 [-105, -116, -154, -214, -252(-263), -312, -350(-361) and -399(-410)]. These results indicate that the repetitious region extends at least 800 bp upstream from the sequenced region.

We have determined partial nucleotide sequences of the fragments next to the sequenced region of Igγ2b=26 as shown in Figure 9. The nucleotide sequence upstream from a Hinf I cleavage site located 1.5 kb away from the 5′ end of the sequenced region (region A in Figure 9) is quite different from the prevalent sequence of the 49 bp repeat unit (Figure 9a). The sequence downstream from the Hinf I site begins to show homology with the prevalent sequence after the first 33 bp, indicating that the 49 bp repeat unit extends about 1.5 kb further toward the 5′ direction from the 5′ end of the sequenced region. Similarly the 3′ terminal border of the repetitive sequence seems to be assigned at about 100 bp from the 3′ end of the sequenced region (Figure 9 c, d and e). The repetitive sequence appears to stretch in the 3.3 kb region between about 1.9 and 5.2 kb 5′ to the 5′ end of the coding region of the γ2b gene.

<af 771>

"Tandem Repetition of 49 bp Unit in $S_{\gamma 1}$ and $S_{\gamma 3}$ Regions"

We have analyzed the nucleotide sequence of Igγ2b=23 to see whether or not similar 49 bp repeating unit is found. We found such sequence between positions -517 and -190 of Igγ2b=23 (Figure 6). The 49 bp unit is repeated six times without an interspaced sequence as shown in Figure 10. Since this portion of Igγ2b=23 is derived from the 5′ flanking region of the γ3 gene, the results indicate that

the $S_{\gamma 3}$ region contains the 49 bp repeating unit. In fact, we have determined the partial nucleotide sequence of the corresponding portion of Igγ3=30 (regions close to two Sac I sites shown in Figure 5) and found that the sequences of the homologous regions of the two clones Igγ2b=23 and Igγ3=30 are identical ~~to each other~~ (data not shown). The repetitive sequence ~~lied~~ lay between ~~about~~ 4.4 and 4.7 kb 5' to the 5' end of the γ3 structural gene (see Figure 5). It is possible, however, that there are additional regions containing 49 bp repeating units in the $S_{\gamma 3}$ region. We will discuss later ~~about~~ the fact that 49 bp repeating unit is not obvious in the $S_{\gamma 3}$ region immediately adjacent to the recombination site.

A similar analysis was done on the nucleotide sequence of the $S_{\gamma 1}$ region previously published (Kataoka et al., 1980). We have also found tandem repetition of 12 repeat units of 49 bp in the region between 2.6 and 3.2 kb upstream from the 5' end of the coding region of the γ1 gene (Figure 10). The repetitious sequence seems to extend further toward the 5' direction in view of the repetitive pattern of the restriction enzyme cleavage sites (A. Shimizu et al., unpublished data).

The nucleotide sequences of the 49 bp units of the $S_{\gamma 1}$ region are similar to one another although the homology is less significant than those of the $S_{\gamma 2b}$ region as shown in Table 1. Likewise, the nucleotide sequences of the 49 bp units of the $S_{\gamma 3}$ region are similar to one another. The nucleotide sequences of the 49 bp units of different S regions are 44 to 51%, ~~in~~ on the average, homologous to each other (Table 1). The prevalent sequences of the 49 bp units of three S regions, $S_{\gamma 1}$, $S_{\gamma 2b}$ and $S_{\gamma 3}$ are compared in Figure 11. It is obvious that three sequences are homologous to each other. These results indicate that the 49 bp units of the S regions of the γ genes have ~~been~~ evolved from a common ancestral sequence. It is worth noting that the recombination sites in the $S_{\gamma 1}$ (MC101 myeloma) and $S_{\gamma 2b}$ (MPC11 myeloma) regions took place at the similar position within the 49 bp unit.

A Portion of the $S_{\gamma 2b}$ Region is deleted from DNAs of γ1 Chain-producing Myelomas

Experiments shown in Figure 3 have indicated that rearrangement has occurred in the 5′ flanking region of the γ2b gene in DNAs of γ1 chain-producing myelomas, MOPC31C and MC101. DNAs from both myelomas contained $S_{\gamma 2b}$ region fragments which were 2.2 kb shorter than those in newborn mouse DNA (Figure 3, lanes d and e). $S_{\gamma 2b}$ region fragment produced by digestion of MOPC31C and MC101 DNAs with Bgl II and Eco RI were also reduced by the same length (2.25 kb) as compared with those of newborn mouse DNA (data not shown). Further analysis clearly demonstrated that such rearrangement in γ1 chain-producing myelomas is due to partial deletion of the repetitious sequence in the $S_{\gamma 2b}$ region. DNAs from newborn mice and various myelomas were cut with Xba I, and their Southern blot filter was hybridized with the repetitious region fragment (fragment A in Figure 9). DNAs of newborn mice and a μ chain-producing myeloma contained fragments (3.3 and 0.8 kb) identical to those of Igγ2b-2206 (Figure 12, lanes a, b and h). In contrast, DNAs of γ1 chain-producing myelomas (lanes c and d) showed no intensely hybridizable bands corresponding to the repetitive region except for a new faint band of 2.5 kb in length. The results indicate that rearrangement of the 5′ flanking region of the γ2b gene in MOPC31C and MC101 DNAs were mediated by deletion of a large part (about 2.25 kb) of the repetitive stretch in the $S_{\gamma 2b}$ region.

Since we detected only one $S_{\gamma 2b}$ gene fragment by digestion with various restriction enzymes, it is likely that MOPC31C and MC101 myeloma DNAs have only one set of the γ2b genes. Although we thought previously that these two myeloma DNAs contained two types of the γ1 genes (Kataoka et al., 1980), we are recently inclined to consider that the germline γ1 gene detected in the myeloma DNAs might be attributable to the contaminant host cell DNA because the relative intensity of the germline gene band varies between batches of DNA. However, we

cannot exclude the possibility that deletion of the $S_{\gamma 2b}$ region in MOPC31C and MC101 occurred in a manner indistinguishable between two chromosomes.

Analyses using other restriction enzymes suggest that the area of deletion ranges from about 1.0 ~ 3.25 kb from the 3′ end of Igγ2b=26. A faint Xba I fragment (2.5 kb) in γ1 chain=producing myelomas (lanes c and d) can be explained by this type of deletion from the restriction map (Figure 2). MPC11 DNA did not contain such fragments as seen in newborn mouse DNA but contained a 3.55 kb band which is coincident to that of Igγ2b=2203 (lanes e and i). The DNA segment of the $S_{\gamma 2b}$ region was completely deleted from γ2a and α chain=producing myeloma DNAs (lanes f and g) in consistence with the deletion model (Honjo and Kataoka, 1978; Coleclough et al., 1980; Cory et al., 1980; Rabbitts et al., 1980; Yaoita et al., 1980a, #1980b). Faint bands of 15, 5.0 and 3.8 kb in length may be ascribed at least in part to cross=hybridization with other S_{γ} regions since we observed that the repetitive region of Igγ2b=2206 cross=hybridized with the 5′ flanking regions of the cloned γ1, γ2a and γ3 gene fragments (N. Takahashi et al., unpublished data; Honjo et al., 198[illegible]).

<af 518>

Discussion

The S region Comprises Repetitive Sequences

The S region was originally defined as the functional region responsible for the class=switch recombination (Kataoka et al., 1980). We have shown that tandem repetitive sequences of regular 49 bp unit are present in the 5′ flanking regions of the γ1, γ2b and γ3 genes where the class=switch recombination takes place. Repetitive sequences are also found in the S_{μ} region (Takahashi et al., 1980; Sakano et al., 1980). Davis et al. (1980b) and Obata et al. (submitted) found that the S_{α} region contains repetition of conserved units of 80 bp. Repetitive sequences of the S regions share short homologous sequences with each

(manuscript submitted for publication 1981)

other (Honjo et al., 1980). These studies have clearly demonstrated that repetitive sequences are characteristic of S regions.

The S region is located a few kb (kilobase pairs) 5′ to the C_H gene. The $S_{\gamma 2b}$ region seems to be present between 1.9 and 5.2 kb 5′ to the γ2b gene (Figure 9). The $S_{\gamma 1}$ (one) and $S_{\gamma 3}$ regions so far sequenced lie at 2.6 ~ 3.2 kb 5′ to the γ1 (one) gene and at about 4.4 ~ 4.7 kb 5′ to the γ3 gene, respectively. We have also found the repetitive region in the 5′ flanking region of the γ2a gene (A. Shimizu et al., manuscript submitted 1980). The S_μ region seems to be located between about 1.3 and 4.8 kb 5′ to the μ gene (Kawakami et al., 1980; Takahashi et al., 1980). The S_α region seems to be present between about 1.2 and 2.4 kb 5′ to the α gene (Davis et al., 1980b). Dunnick et al. (1980) found 49 bp repeating units in the intervening sequence separating a V_H and C_H (γ1) (one) genes (one) of a myeloma mutant (IF2) which has lost the complete C_{H1} (one) domain of the γ1 gene. Although the location of their repetitive sequences in the germ line gene is not known, the sequences of their repeating units are homologous to those of our S_γ region sequences shown in Figures 7 and 10. We presume that their repetitive region is located at the 5′ side of the $S_{\gamma 1}$ (one) region sequenced in the present study.

2 of 771

What Role Do Repetitive Sequences Play in the Recognition of Class-Switch Recombination?

Comparison of the nucleotide sequences of the immediate proximity of the recombination sites published previously (Kataoka et al., 1980; Takahashi et al., 1980) and determined in the present study does not reveal a unique sequence nor a unique structure such as that was shown in the V-J recombination site (Max et al, 1979; Sakano et al., 1979 and 1980; Seidman et al., 1979; Early et al., 1980; Honjo et al., 1980).

It is known that the tandem repetitive sequence increases the chance of the homologous recombination (Russell et al., 1970). Since the repeat units of the

different S_{γ} regions contain (albeit not identical, similar sequences,) it is likely that the homologous recombination takes place at a considerable rate between different S_{γ} regions, resulting in the H chain class switch from γ3 to γ1 or to γ2b. Southern blot hybridization experiments showed that the S_{μ} region had extensive homology with the S_{α} region and less but significant homology with the $S_{\gamma 1}$, $S_{\gamma 2a}$, $S_{\gamma 2b}$ and $S_{\gamma 3}$ regions (Honjo et al., 1980; N. Takahashi et al., unpublished data). We have not completed the nucleotide sequence determination of the S_{μ} and S_{α} regions. A partial sequence of this region clearly indicates the presence of repetitive sequences (Takahashi et al., 1980; Sakano et al., 1980). The S_{μ} region shares short common sequences such as AGCT, TGGG and AGCTGGGG with S_{γ} and S_{α} regions.

Although the $S_{\gamma 3}$ region immediately adjacent to the recombination site of MPC11 does not contain the characteristic repeating unit of 49 bp, the nucleotide sequence in that region share short common sequences such as GAGCTG, GAGC, GAGCT and TGGGG with that of the $S_{\gamma 2b}$ region. The results may implicate that it is abundant short common sequences rather than regular 49 bp repeating unit that are important for S=S recombination.

We have proposed a model that class=switch recombination is mediated by an unequal crossing=over event between sister chromatids (Honjo et al., 1980). It may be possible that decades of short common sequences as shown above are enough to make a transient synapsis of two DNAs when two sister chromatids are brought close to each other. Such interaction may be helped by a specific protein(s) which recognizes some short sequences. Such protein per se or collaboratory proteins may catalize excision and ligation of DNAs with recognition of similar short sequences. If this is the case, actual switch recombination sites are widely spread in and around the S region. We have determined one switch site in the $S_{\gamma 3}$ region, one in the $S_{\gamma 1}$ region and two in the $S_{\gamma 2b}$ region. The switch

site of MOPC141 γ2b gene is about 1.7 kb 3′ to that of MPC11 γ2b gene (Takahashi et al., 1980). The common tetranucleotide $5'\frac{!}{N}AGNTG\frac{!}{N}3'$ is found in the immediate proximity of these two switch sites in the $S_{\gamma 2b}$ region, the middle nucleotide (N) being the recombination site in both cases (Takahashi et al., 1980; Honjo et al., 1980). It may also worth noting that inverted repeat sequences are present $45\frac{!}{N}74$ bases 5′ to the recombination site in Igγ2b=23 ($S_{\gamma 3}$ region) and $2\frac{!}{N}25$ bases 3′ to the recombination site in Igγ2b=26 ($S_{\gamma 2b}$ region) as shown in Figure 13. Inverted repeat sequences are present in the similar position of many 49 bp repeating units of the $S_{\gamma 2b}$ region.

A combined structure of tandem repetition and short recognition sequence will provide a large number of the possible recombination sites in the S region and increase the chance of the class=switch recombination. As pointed out before, the class=switch recombination does not have to be highly specific to the nucleotide joined together (Takahashi et al., 1980). Instead, the class=switch recombination is expected to be efficient since it takes place during relatively short period of time after stimulation with an antigen. The above model satisfies these biological features required for the class=switch recombination.

<af 771>

Deletion in the S region

We have shown that a portion of the $S_{\gamma 2b}$ region is deleted from DNAs of the γ1 chain=producing myelomas. The deletion seems to have taken place at the identical location in MOPC31C and MC101 myelomas which have different types of rearrangements in the γ1 gene fragment (Kataoka et al., 1980). Several γ2b chain=producing myelomas (MOPC141, MPC11 and BKC F1 #15) were shown to have rearrangements in the 3′ flanking region of the γ2b gene, which is actually close to the $S_{\gamma 2a}$ region since the γ2a gene is located only 15 kb 3′ to the γ2b gene (Honjo et al., 1980; A. Shimizu et al., 1980).

In summary, there are a number of myelomas in which the 3′ side of the

expressed $C_H^{\wedge}$ gene is rearranged. Some of them are clearly explained by the deletion of a portion of the S region. It is difficult to envisage a direct cause and result relationship between the expression of a $C_H^{\wedge}$ gene and rearrangement of the 3' flanking region of the expressed $C_H^{\wedge}$ gene. It is possible that a part of the S region tends to be deleted by itself without involving class-switch recombination. It may be relevant to this point that other strains of mice have the $S_\mu^{\wedge}$ region of varied length such that C57BL has the $S_\mu^{\wedge}$ region 2 kb longer than that of BALB/c (Y. Yaoita and T. Honjo, unpublished data).

It is always important to keep in mind that these studies done on myeloma cells may contain secondary rearrangements induced during propagation of myeloma cells. Analysis of normal B lymphocytes is necessary to clear this point.

<af 772>

Experimental Procedures

Materials

γ-$^{32}P_N^1$ATP (spec. act. 5500 Ci/mmole) and α-$^{32}P_N^1$dCTP (spec. act. 2000 Ci/mmole) were obtained from New England Nuclear (Boston, Massachusetts). Sources of restriction enzymes, polynucleotide kinase and other enzymes were described previously (Kataoka et al., 1979; Honjo et al., 1979). Bacterial strains used for in vitro packaging were kindly supplied by Dr. F. Blattner of the University of Wisconsin. Myeloma tumor cell lines were supplied by Dr. M. Potter of NIH, Dr. S. Migita of Kanazawa University and Dr. P. Leder of NIH and were propagated by subcutaneous injection as described (Sibinovic et al., 1976).

<af 773>

Cloning Procedure

High molecular weight DNA was prepared from MPC11 (eleven) myeloma tumor or newborn mice, digested partially with Eco RI, and fractionated by agarose gel (0.7%) electrophoresis as described (Polsky et al., 1978). The partially enriched immu-

noglobulin gene fragments were ligated with λgtWES (Leder et al., 1977) or Charon 4A (Blattner et al., 1977) outer fragments, packaged in vitro into phage coats (Blattner et al., 1979) and the resultant plaques were screened as described (Benton and Davis, 1977) using the nick-translated Hha I fragment of cloned γ2b cDNA (pG2b-4) as a probe (Kataoka et al., 1979). The recombinant phage DNAs were cleaved with Eco RI and the inserts were recloned into pBR322 (Bolivar et al., 1977). Recombinant phage and plasmid DNAs were prepared as described by Tiemeier et al. (1977) and Clewell and Helinski (1969), respectively. All the cloning experiments were carried out in a P2 facility using EK2 host-vector systems.

<af 773>

Southern Blot Hybridization

DNA (10 μg) was digested with appropriate amounts of various restriction enzymes at their optimal conditions, electrophoresed in agarose gels (Type I, Sigma Chemical Co., St. Louis, Missouri), and blotted to nitrocellulose filters (Schleicher and Schuell) as described (Southern, 1975). Nick translation of DNA fragments was performed using α-^{32}P-dCTP (2000 Ci/mmole) and E. coli DNA polymerase I (Boehringer) as described (Maniatis et al., 1975). The Southern blot filters were hybridized with appropriate probes as described previously (Kataoka et al., 1979; Honjo et al., 1979).

<af 773>

DNA Sequencing

5'-terminal labelling of restriction fragments and nucleotide sequence determination was performed according to the method of Maxam and Gilbert (1977) with minor modifications described previously (Honjo et al., 1979).

<af 739>

Acknowledgement

We thank Dr. D. Baltimore for a 73 cDNA clone and Drs. Y. Suzuki and S. Nakanishi for critical reading of the manuscript. This investigation was supported by grants from Mitsuhisa Memorial Cancer Research Fund, Asahi Fund for Scientific Research, and Ministry of Education, Science and Culture of Japan.

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. Section 1734 solely to indicate this fact.

September 9, 1980 October 29, 1980

References

Abney, E.R., Cooper, M.D., Kearney, J.F., Lawton, A.R. and Parkhouse, R.M.E. (1978) J. Immunol. 120, 2041-2049.

Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.

Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston=Thompson, K., Faber, H.E., Richard, J.E., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977) Science 196, 161-169.

Blattner, F.R., Blechl, A.E., Denniston=Thompson, K., Faber, H.E., Richard, J.E., Slighton, J.L., Tucker, P.W. and Smithies, O. (1978) Science 202, 1279-1283.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) Gene 2, 95-113.

Clewell, D. and Helinski, D. (1969) Proc. Natl. Acad. Sci. USA 62, 1159-1166.

Coleclough, C., Cooper, D. and Perry, R.P. (1980) Proc. Natl. Acad. Sci. USA 77, 1422-1426.

Cory, S. and Adams, J.M. (1980) Cell 19, 37-51.

Cory, S., Jackson, J. and Adams, J.M. (1980) Nature 285, 450-456.

Davis, M.M., Calame, K., Early, P.W., Livant, D.L., Joho, R., Weissman, I.L. and Hood, L. (1980a) Nature 283, 733-739.

Davis, M.M., Kim, S.K. and Hood, L.E. (1980b) Science 209, 1360-1365.

Dunnick, W., Rabbitts, T.H. and Milstein, C. (1980) Nature 286, 669-675.

Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980) Cell 19, 981-992.

Fu, S.M., Winchester, R.J. and Kunkel, H.G. (1975) J. Immunol. 114, 250-252.

Gearhart, P.J., Sigal, N. and Klinman, N.R. (1975) Proc. Natl. Acad. Sci. USA 72, 1707-1711.

Goding, J.W. and Layton, J.E. (1976) J. Exp. Med. 144, 852-857.

Honjo, T. and Kataoka, T. (1978) Proc. Natl. Acad. Sci. USA 75, 2140-2144.

Honjo, T., Obata, M., Yamawaki=Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N. and Mano, Y. (1979) Cell 18, 559-568.

Q pdate Honjo, T., Kataoka, T., Yaoita, Y., Shimizu, A., Takahashi, N., Yamawaki=Kataoka, Y., Nikaido, T., Nakai, S., Obata, M., Kawakami, T. and Nishida, Y. (198?) Cold Spring Harbor Symp. Quant. Biol. 45 in press.

Kataoka, T., Yamawaki=Kataoka, Y. Yamagishi, H. and Honjo, T. (1979) Proc. Natl. Acad. Sci. USA 76, 4240-4244.

Kataoka, T., Kawakami, T., Takahashi, N. and Honjo, T. (1980) Proc. Natl. Acad. Sci. USA 77, 919-923.

date Q Kawakami, T., Takahashi, N. and Honjo, T. (1980) Nucl. Acids Res. 8, 3933-3945 ~~in press.~~

Leder, P., Tiemeier, D. and Enquist, L. (1977) Science 196, 175-177.

Maki, R. Traunecker, A., Sakano, H., Roeder, W. and Tonegawa, S. (1980) Proc. Natl. Acad. Sci. USA 77, 2138-2142.

Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Proc. Natl. Acad. Sci. USA 72, 1184-1188.

Max, E.E., Seidman, J.G. and Leder, P. (1979) Proc. Natl. Acad. Sci. USA 76, 3450-3454.

Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.

Nossal, G.J.V., Warner, N. and Lewis, H. (1971) Cell. Immunol. 2, 41-53.

Polsky, F., Edgell, M.H., Seidman, J.G. and Leder, P. (1978) Anal. Biochem. 87, 397-410.

Rabbitts, T.H., Forster, A., Dunnick, W. and Bentley, D.L. (1980) Nature 283, 351-356.

Rao, D.N., Rudikoff, S. Krutzsch, H. and Potter, M. (1979) Proc. Natl. Acad. Sci. USA 76, 2890-2894.

Russell, R.L., Abelson, J.N., Landy, A., Gefter, M.L., Brenner, S. and Smith, J.D.

Obata, M., Kataoka, T., Nakai, S., Yamagishi, H., Takahashi, N., Yamawaki-Kataoka, Y., Nikaido, T., Shimizu, A. and Honjo, T. (1981). Proc. Nat. Acad. Sci. USA in press.

(1970) J. Mol. Biol. 47, 1-13.

Sakano, H., Huppi, K., Heinrich, G. and Tonegawa, S. (1979) Nature 280, 288-294.

Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. and Tonegawa, S. (1980) Nature 286, 676-683.

Seidman, J.G., Max, E.E. and Leder, P. (1979) Nature 280, 370-375.

Sibinovic, K.H. and Potter M. (1973) In Catalogue of Murine Plasma Cell Tumors, 1st ed. (Litton Bionetics, Kensington, Maryland, USA) Apendix 1-3. pp 000-000.

Sledge, C., Fair, D.S., Black, B., Krueger, R.G. and Hood, L. (1976) Proc. Natl. Acad. Sci. USA 73, 923-927.

Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.

Takahashi, N., Kataoka, T. and Honjo, T. (1980) Gene ~~in press.~~ 11, 117-127

Tiemeier, D.C., Tilghman, S.M. and Leder, P. (1977) Gene 2, 173-191.

Tucker, P.W., Marcu, K.B., Newell, N., Richards, J. and Blattner, F.R. (1979) Science 206, 1303-1306.

Wang, A.=C., Wang, I.Y. and Fudenberg, H.H. (1977) J. Biol. Chem. 252, 7192 1/N C 7199.

Yamawaki=Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M. and Honjo, T. (1980) Nature 283, 786-789.

Yaoita, Y. and Honjo, T. (1980a) Biomedical Res. 1, 164-175.

Yaoita, Y. and Honjo, T. (1980b) Nature ~~in press~~. 286, 850-853.

Shimizu, A., Takahashi, N., Yamawaki-Kataoka, Y., Nishida, Y., Kataoka, T. and Honjo, T. (1980) Nature, in press.

<af 52 >

03024

Table 1. Homology of Repeat Units

Repeat Units Compared (one)	% Average Homology (standard deviation)
$S_{\gamma 1}/S_{\gamma 1}$	46.4 (11.3)
$S_{\gamma 1}/S_{\gamma 2b}$	49.4 (9.7)
$S_{\gamma 1}/S_{\gamma 3}$	43.6 (9.2)
$S_{\gamma 2b}/S_{\gamma 2b}$	73.2 (7.9)
$S_{\gamma 2b}/S_{\gamma 3}$	50.6 (6.0)
$S_{\gamma 3}/S_{\gamma 3}$	44.3 (7.5)

Grk. l.c. gamma

A pair of repeat units of the S regions indicated was compared with all the possible combinations. The repeat unit sequences were taken from Figure 7 and 10.

Legends to Figures

<af 520>

Figure 1. A Model for Heavy Chain Class Switch

Recombination events which occur to form a complete γ2b gene during differentiation of a B lymphocyte are schematically represented. The model was taken from Kataoka et al. (1980) with slight modifications. The presence of D segments was indicated by recent studies from several laboratories (Early et al., 1980; Honjo et al., 1980; Sakano et al., 1980).

<af 521>

Figure 2. Comparative Restriction Cleavage Maps of the Germline and the Rearranged γ2b Chain Genes and Sequencing Strategies

Various restriction enzyme cleavage sites of cloned γ2b gene segments from germline DNA (Igγ2b=2206) or from MPC11 DNA (Igγ2b=2203) are represented with the direction of transcription from left to right. Closed squares in the bars represent the γ2b gene structural sequence, which is interrupted by three intervening sequences (Kataoka et al., 1979; Tucker et al., 1979; Yamawaki-Kataoka et al., 1980). The third and forth horizontal bars are enlargements of a region 2.0 ~~Eo~~ 3.6 kb 5′ to the structural γ2b gene of Igγ2b=2206 and Igγ2b=2203, respectively. Horizontal arrows indicate the directions and ranges of sequencing. Black squares A and B indicate a Bam HI fragment used as a probe for the experiment shown in Figure 3 and 1.7 kb Hinc II=Xba I fragment used in Figure 8, respectively.

<af 571>

Figure 3. Hybridization of Newborn Mouse and Various Myeloma DNAs with a γ2b Gene Fragment

Ten μg each of DNA of newborn mouse (lane a), MOPC104E (μ chain producer, lane b), J606 (γ3, lane c), MOPC31C (γ1, lane d), MC101 (γ1, lane e), MOPC141 (γ2b, lane f), MPC11 (γ2b, lane g), RPC5 (γ2a, lane h), and MOPC315 (α, lane i) were cut with restriction endonucleases indicated and electrophoresed in 0.7%

agarose gels. Approximately 100 pg of λgt WES ∪ Igγ2b=2206 (lane j) and λgt WES ∪ Igγ2b=2203 (lane k) were cut with the same enzymes indicated and electrophoresed in the same gel. Southern blots of the gels were hybridized with the nick translated (spec. act. ≈ 2 x 10^8 cpm/μg DNA) fragment A of pIgγ2b=22 shown in Figure 2. Molecular sizes of the bands were calculated with Hind III digests of λcI857 DNA as molecular markers."

<af 52φ>

Figure 4." Hybridization of Newborn Mouse and MPC11 DNAs with the Rearranged γ2b Gene Fragment and of Cloned 18 kb DNA with Cloned γ3 Chain cDNA"

(A) 10 μg each of newborn mouse and MPC11 DNAs were digested with Eco RI, electrophoresed in a 0.5% agarose gel and blotted to a nitrocellulose filter. The blot was hibridized with pIgγ2b=23 DNA labeled by nick translation (spec. act. ≈ 1 x 10^8 cpm/μg) and autoradiographed.

(B) 0.5 μg each of Charon 4A ∪ Igγ3=30 DNA was cut with the restriction enzymes indicated and electrophoresed in a 0.5% agarose gel. The Southern blot of the gel was hybridized with cloned γ3 chain cDNA labeled with ^{32}P by nick translation (spec. act. ≈ 5 x 10^7 cpm/μg)."

<af 52φ>

Figure 5." Comparative Restriction Maps of Germline γ3 Gene and Igγ2b=23"

Insert DNA segments of Charon 4A ∪ Igγ3=30 and pIgγ2b=23 are represented by upper and lower horizontal bars, respectively, with the direction of transcription from left to right. Closed square with serrated borders indicates the approximate location of γ3 structural gene sequence determined by Southern blot hybridization (Figure 4B) and R loop analysis (S. Nakai, Y. Yamawaki-Kataoka and T. Honjo, unpublished data). Vertical broken line indicates the recombination site in Igγ2b=23. Closed boxes at Sac I sites of Igγ3=30 indicate the region sequenced."

<af 521>

Figure 6." Nucleotide Sequence Surrounding the Recombination Site of Igγ2b=26 Igγ2b=23"

The nucleotide sequences of a region located between two Xba I cleavage sites of 4.4 and 6.1 kb from the 5' terminus of Igγ2b=26 and the corresponding region between a Sac I site and a Xba I site 4.65 and 6.1 kb from the 5' terminus of Igγ2b=23 (see Figure 2) are presented with the direction of transcription of the structural gene. An arrow head indicates the recombination site. O in position 696 indicates that the base sequence was not determined."

<af sa1>

Figure 7. "Tandem Repetitious Units of 49 bp in the $S_{\gamma 2b}$ Region"

"Nucleotide sequence of Igγ2b=26 shown in Figure 6 is aligned in terms of six repeating groups of 245 or 295 bp, each of which comprises five or six repeat units of 49 bp. Homologous bases are boxed. The prevalent sequence of the 49 bp unit is shown at the bottom. Base positions shown at right are identical to those in Figure 6."

<af sa2>

Figure 8. "Partial Digestion of the Igγ2b=26 Fragment Flanking the 5' End of the Sequenced Region"

"The 3.3 kb long Xba I fragment of pIgγ2b=26 was terminally labeled by polynucleotide kinase and [γ-^{32}P]ATP, and cut with Hinc II. The resultant 1.7 kb long Hinc II=Xba I fragment (fragment B in Figure 2) labeled only at the Xba I site was partially digested with Ava II ~~(A)~~ or Alu I ~~(B)~~ electrophoresed in 5% polyacrylamide gels and autoradiographed. Approximately 2 μg of the [^{32}P]=fragment B was cleaved partially with Ava II or Alu I. The amount of enzymes and incubation time are as follows. Lanes a, b, c, d and e, one unit of Ava II for 2, 10, 20, 40 and 80 min, respectively; lane f, 4 units of Ava II for 10 min. Lanes g, h and i, 0.8 unit of Alu I for 2, 10 and 40 min, respectively; lanes j, k and l, 4 units of Alu I for 10, 20 and 40 min, respectively. The molecular sizes of the bands were calculated with Hae III digests of colicin E1 DNA as a molecular size standard and represented in base pairs."

Figure 1

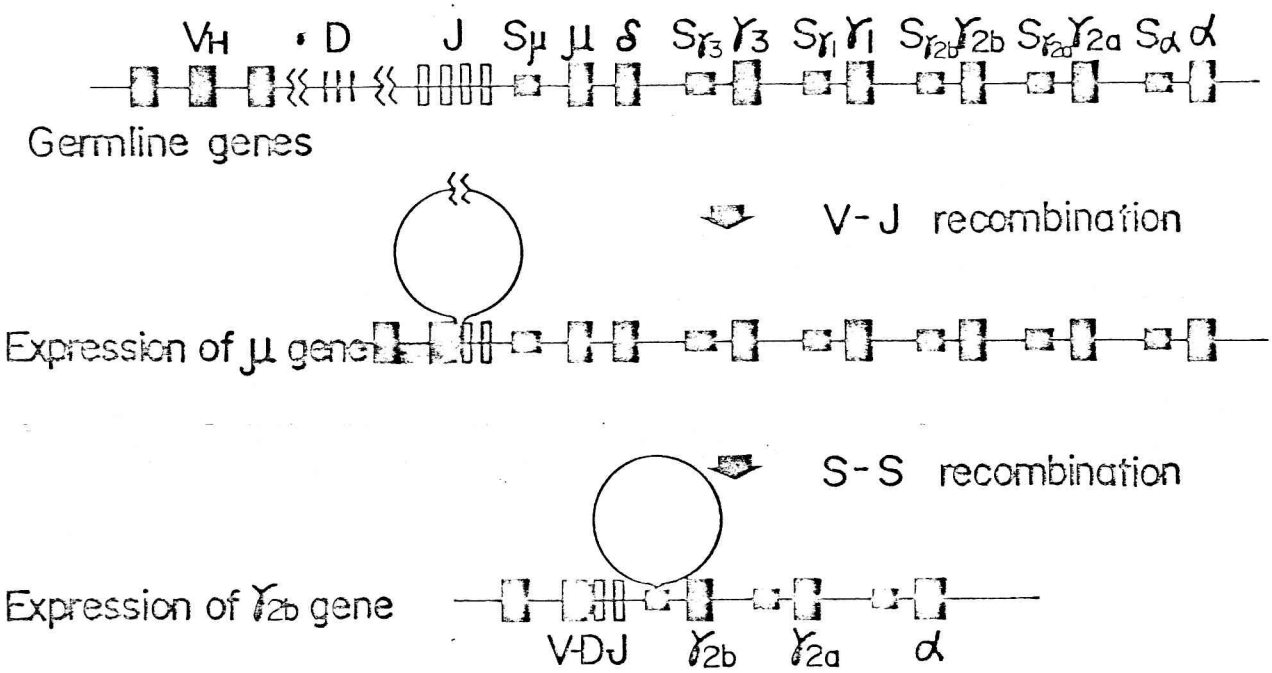

VH
D
J
Sμ
μ
δ
Sγ3
γ3
Sγ1
γ1
Sγ2b
γ2b
Sγ2a
γ2a
Sα
α
Germline genes
V-J recombination
Expression of μ gene
S-S recombination
Expression of γ2b gene
V-DJ
γ2b
γ2a
α

Figure 2

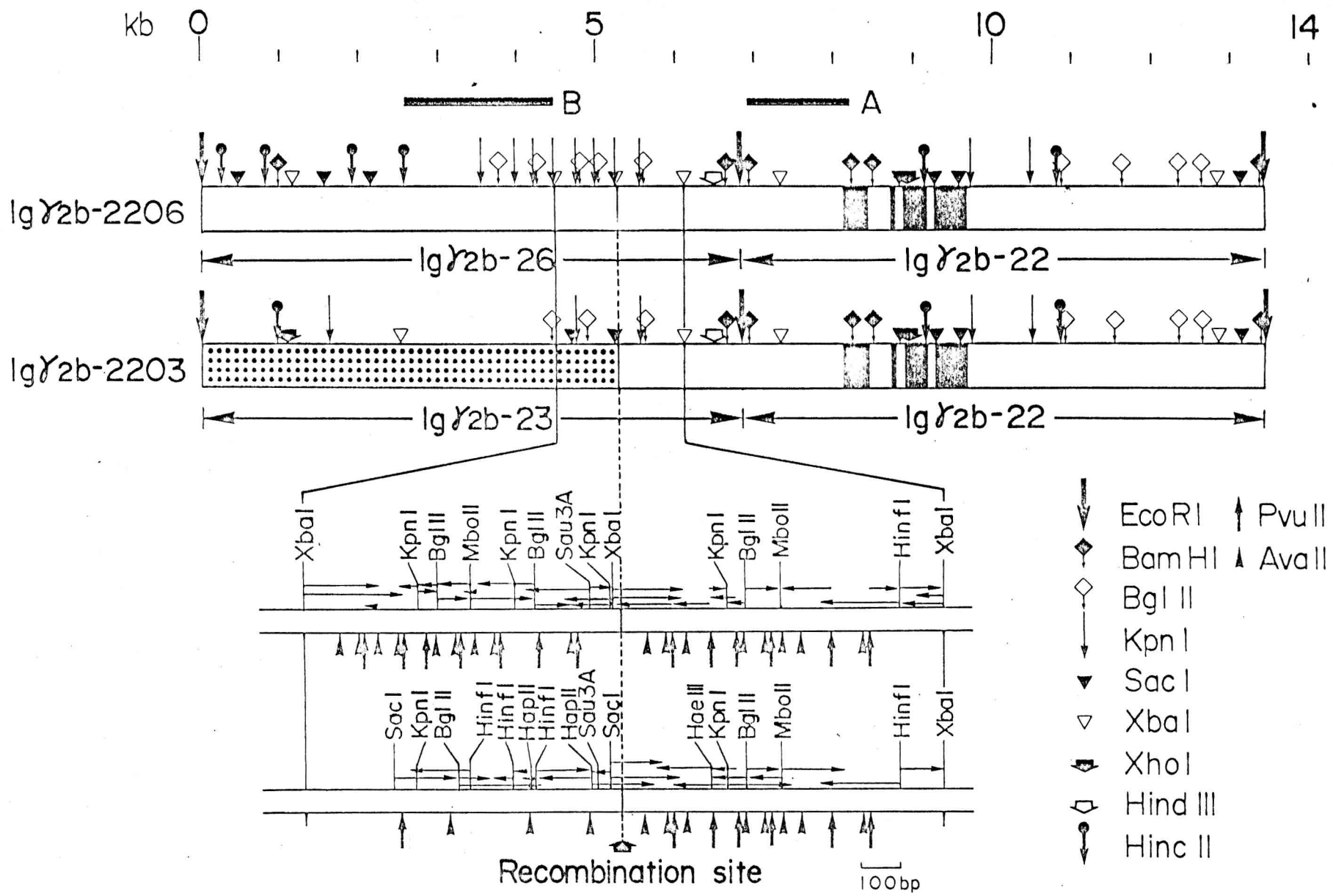

kb
0
5
10
14
B
A
Ig γ2b-2206
Ig γ2b-26
Ig γ2b-22
Ig γ2b-2203
Ig γ2b-23
Ig γ2b-22
Xba I
Kpn I
Bgl II
Mbo II
Kpn I
Bgl II
Sau3A
Kpn I
Xba I
Kpn I
Bgl II
Mbo II
Hinf I
Xba I
Sac I
Kpn I
Bgl II
Hinf I
Hinf I
Hap II
Hinf I
Hap II
Sau3A
Sac I
Hae III
Kpn I
Bgl II
Mbo II
Hinf I
Xba I
Recombination site
100bp
EcoR I
Pvu II
Bam HI
Ava II
Bgl II
Kpn I
Sac I
Xba I
Xho I
Hind III
Hinc II

Figure 3

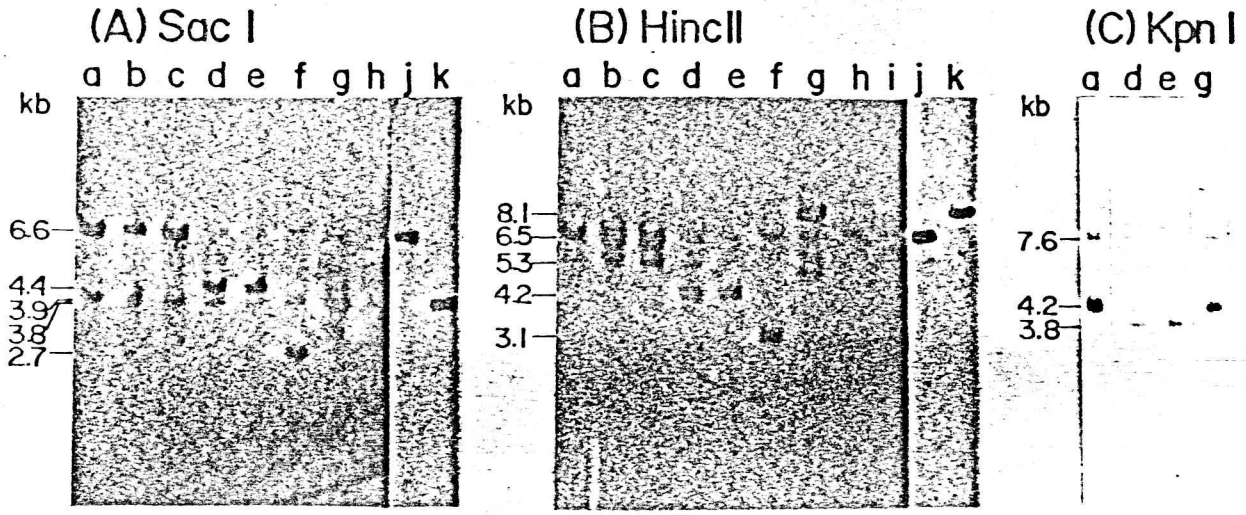

(A) Sac I
a b c d e f g h j k
kb
6.6
4.4
3.9
3.8
2.7
(B) HincII
a b c d e f g h i j k
kb
8.1
6.5
5.3
4.2
3.1
(C) Kpn I
a d e g
kb
7.6
4.2
3.8

Figure 4

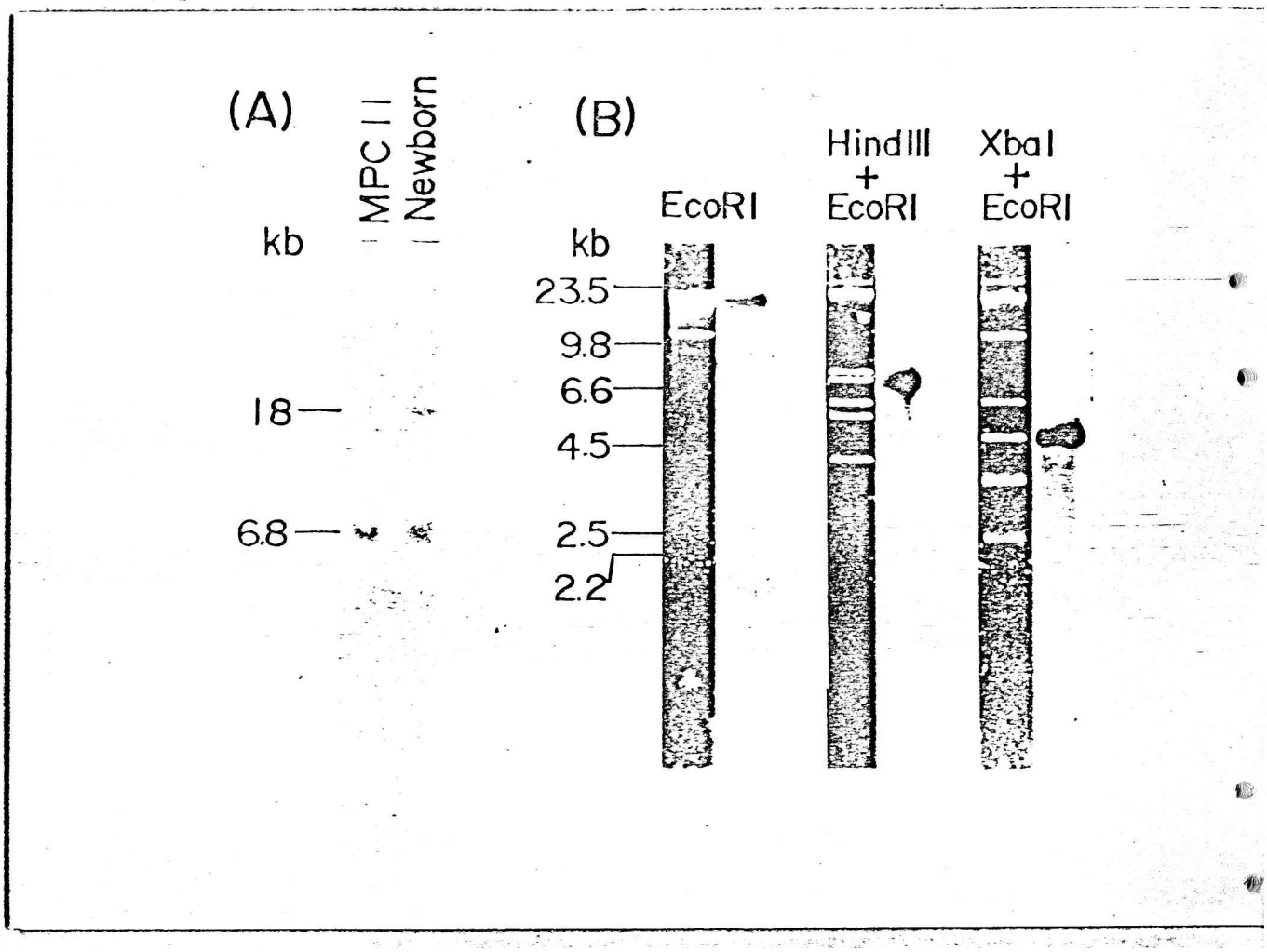

(A)
MPC 11
Newborn
kb
18
6.8
(B)
EcoRI
HindIII + EcoRI
XbaI + EcoRI
kb
23.5
9.8
6.6
4.5
2.5
2.2

Figure 5

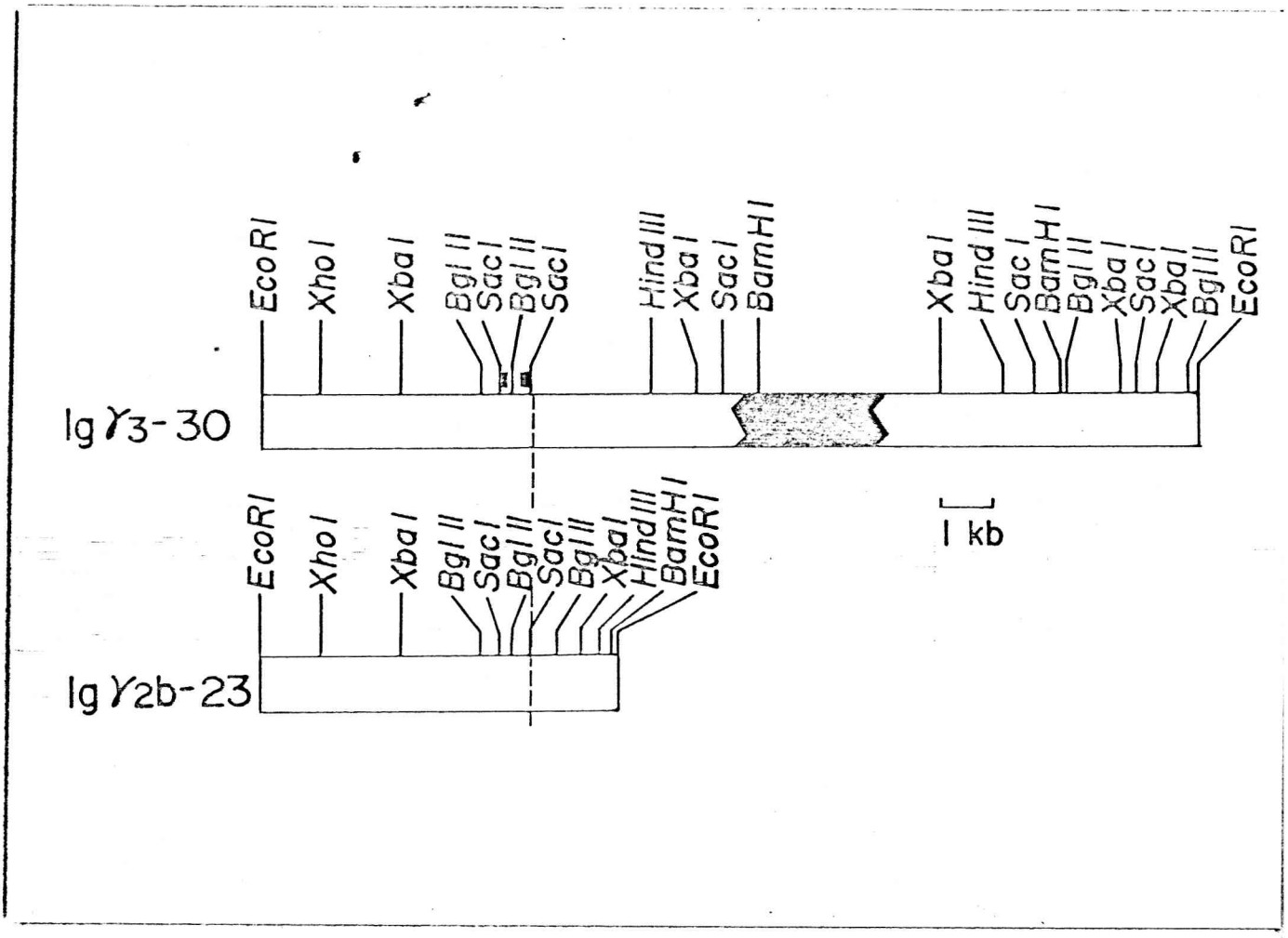

EcoRI
Xho I
Xba I
Bgl II
Sac I
Bgl II
Sac I
Hind III
Xba I
Sac I
BamH I
Xba I
Hind III
Sac I
BamH I
Bgl II
Xba I
Sac I
Xba I
Bgl II
EcoRI
lg γ3-30
1 kb
EcoRI
Xho I
Xba I
Bgl II
Sac I
Bgl II
Sac I
Bgl II
Xba I
Hind III
BamH I
EcoRI
lg γ2b-23

Figure 6

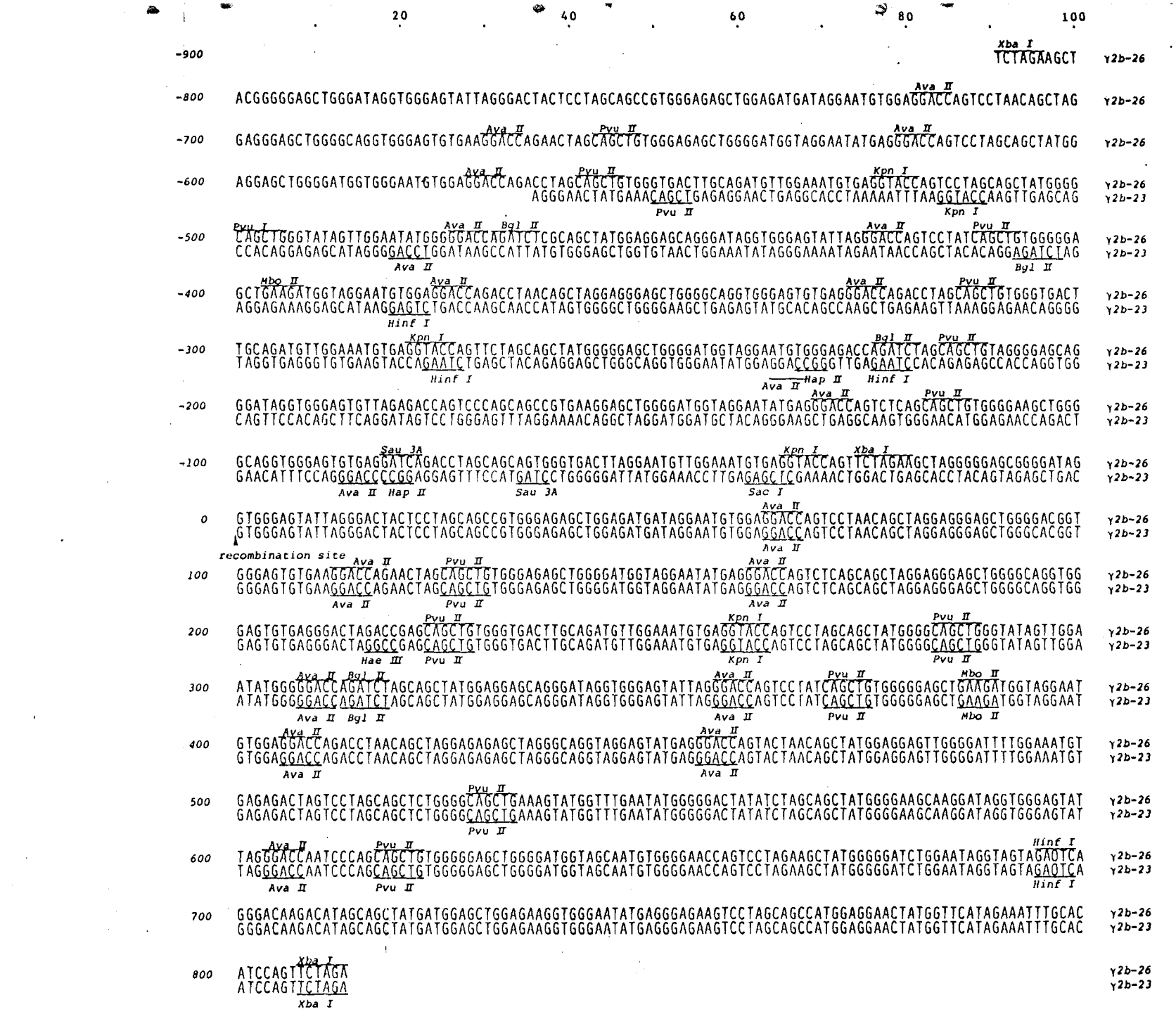

20
40
60
80
100
-900
-800
-700
-600
-500
-400
-300
-200
-100
0
100
200
300
400
500
600
700
800
recombination site
Xba I
Ava II
Pvu II
Kpn I
Bgl II
Mbo II
Hinf I
Sau 3A
Hap II
Sac I
Hae III
γ2b-26
γ2b-23

Figure 7

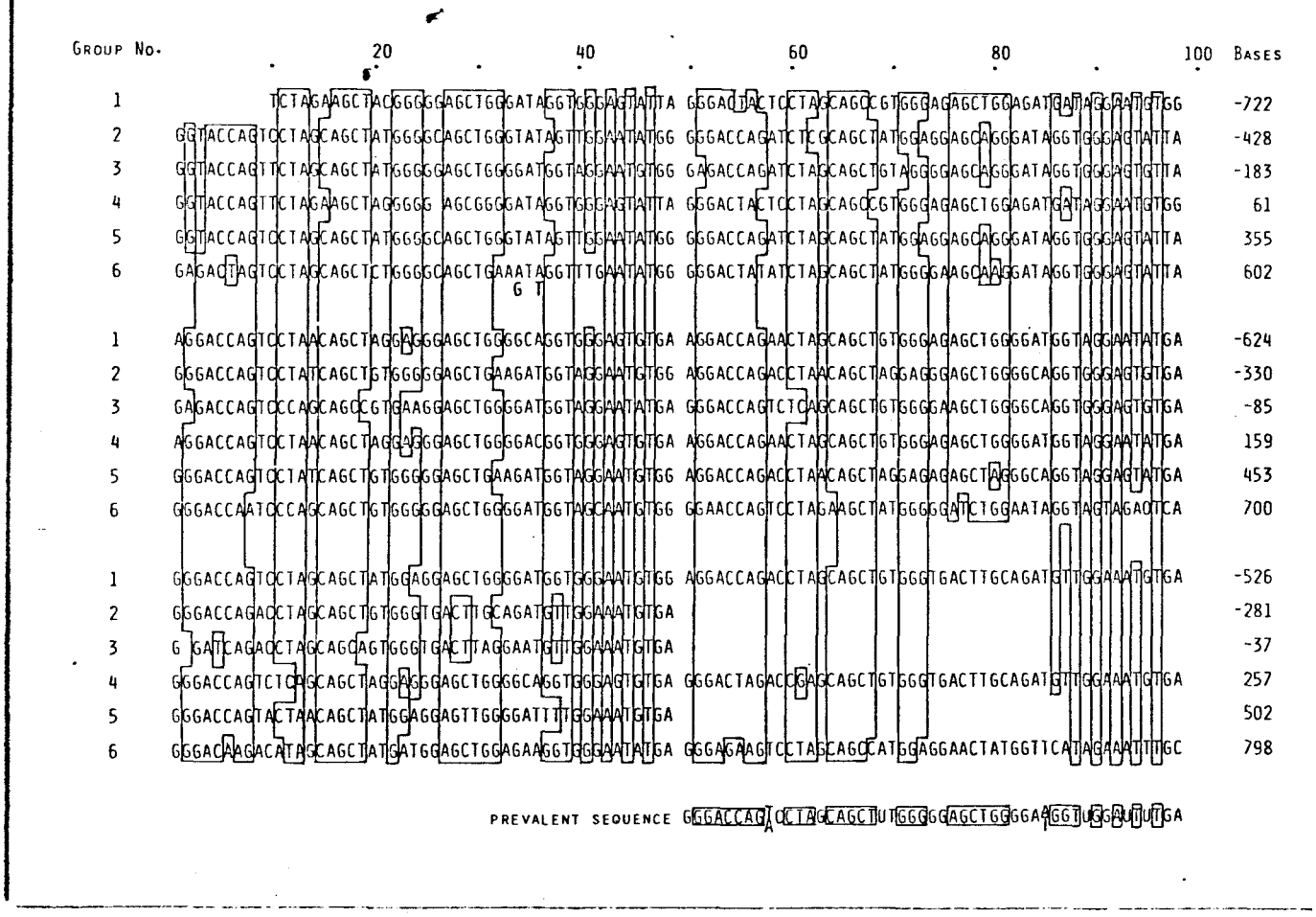

GROUP No.
20
40
60
80
100
BASES
1 -722
2 -428
3 -183
4 61
5 355
6 602
1 -624
2 -330
3 -85
4 159
5 453
6 700
1 -526
2 -281
3 -37
4 257
5 502
6 798
PREVALENT SEQUENCE

Figure 8

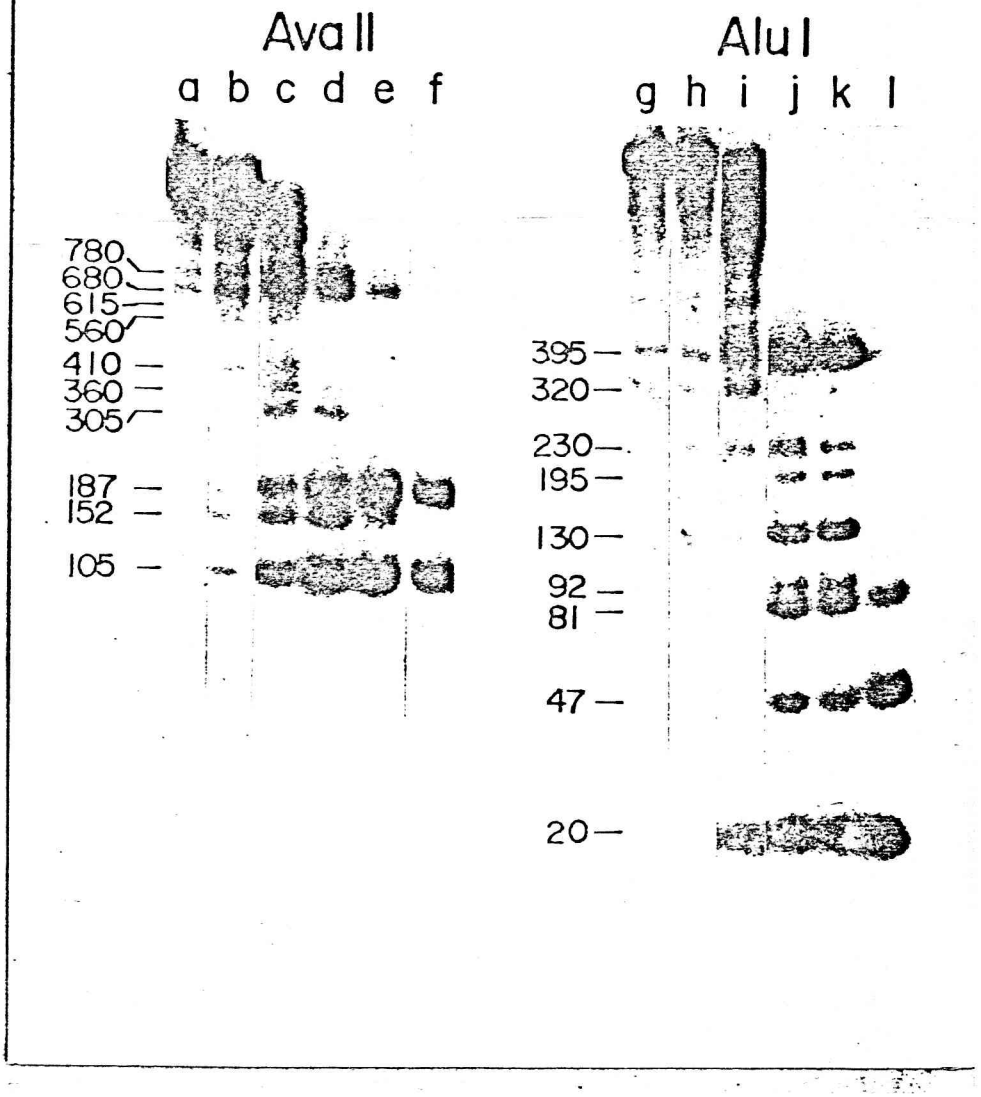
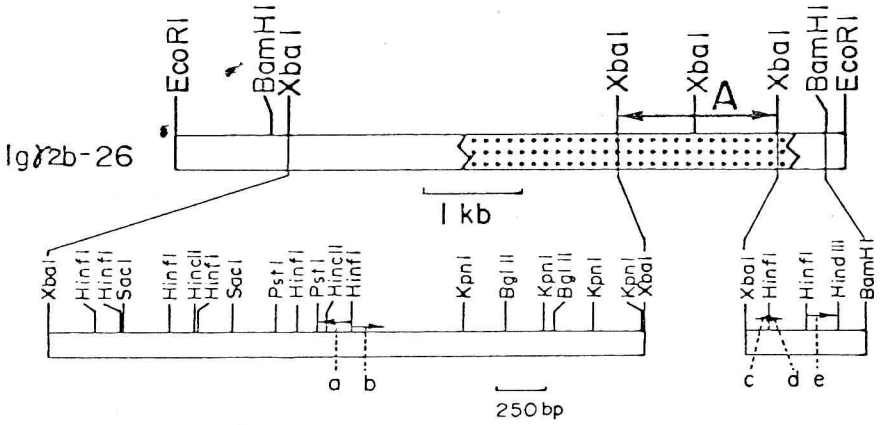

AvaII
a b c d e f
780
680
615
560
410
360
305
187
152
105
AluI
g h i j k l
395
320
230
195
130
92
81
47
20

Figure 9



$S_{\gamma 2b}$ prevalent sequence GGGACCAG$^{T}_{A}$CCTAGCAGCTPTGGGGGAGCTGGGGA$^{A}_{T}$GGTPGGAPTPTGA

a GGGAGCAGGTGTGGTTAACAGGGGTCCAAGGCACCTTTCAGGAGGTAGC
AAAAAACTTTACAGGAGTTCTGGAACAGAAAGGAGTTGGGAAGGCAAGC
TGAGCATCTGTAAAGATACTGTGGCAGTTGTCGGAGTC

b GAGTCTGGGTCTCAGAGCTACCCTAGGGACTAA
GGGACAAATCCGAGTAGCCACGGGGGAGTGGGGGATGGTAGGAATGTAA
GGGATTAGTCCTAGCAGTTATGGGGAAGCTATGAATGGTAGGAATGTTG

c CTTCTAGCAGCTATGGGG

d TCTTGGTCAGGTGAGGGTGATTTTAC

e TCAGTAAACTTTGGAGTAT
GGTGAGCATCCAGAAAATCTGTGGTGCAACTGTAACAGGTGGGGTGTGA
GGAACCAGACCCAGTAAGAATGGGATACTTGGGGAGAGCGGCATGTGGA
GGGACCACCAGTGGTTAAGGA

Figure 10

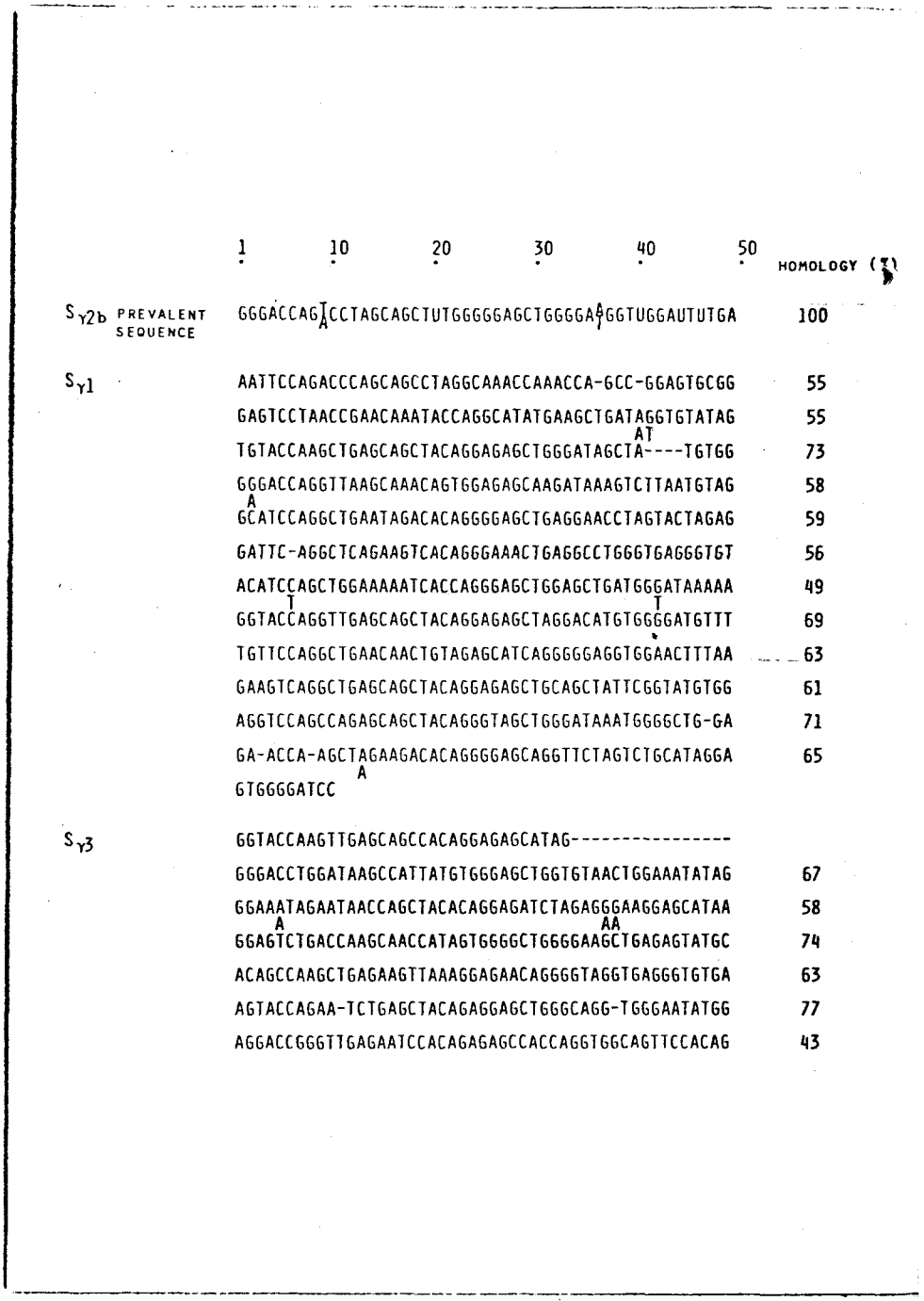
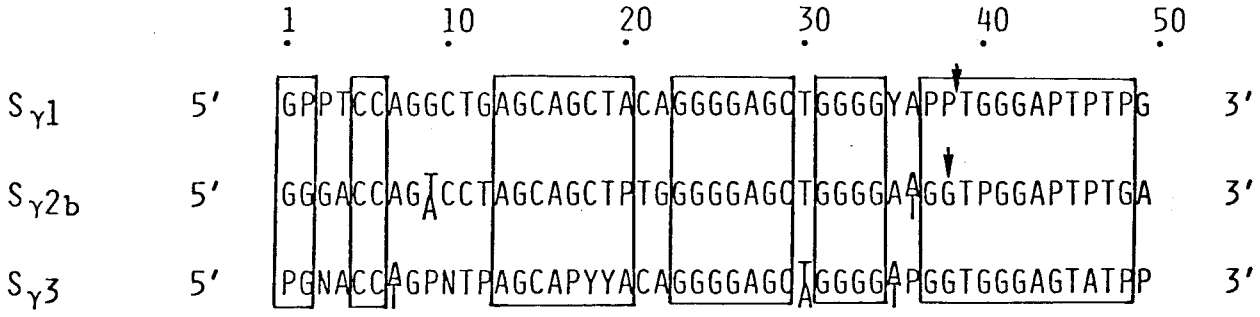
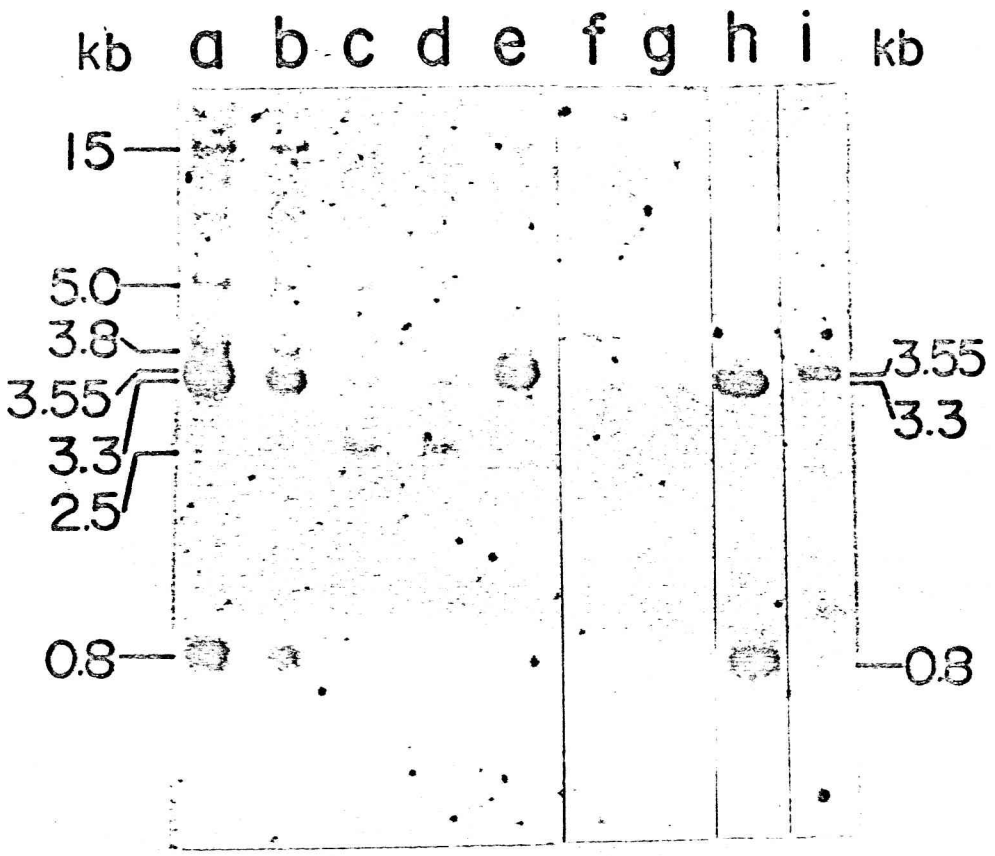
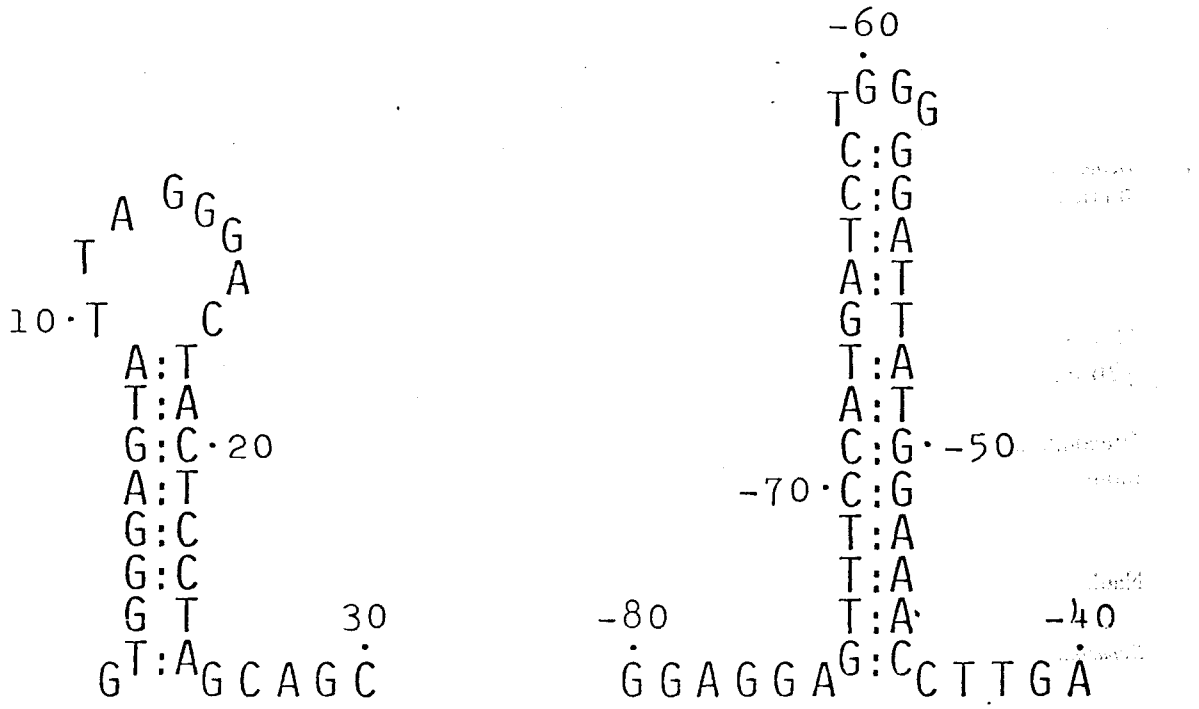

1
10
20
30
40
50
HOMOLOGY (%)
$S_{\gamma 2b}$ PREVALENT SEQUENCE
GGGACCAGTCCTAGCAGCTTTGGGGGAGCTGGGGATGGTTGGATTTTGA
100
$S_{\gamma 1}$
AATTCCAGACCCAGCAGCCTAGGCAAACCAAACCA-GCC-GGAGTGCGG
55
GAGTCCTAACCGAACAAATACCAGGCATATGAAGCTGATAGGTGTATAG
55
TGTACCAAGCTGAGCAGCTACAGGAGAGCTGGGATAGCTA----TGTGG
73
GGGACCAGGTTAAGCAAACAGTGGAGAGCAAGATAAAGTCTTAATGTAG
58
GCATCCAGGCTGAATAGACACAGGGGAGCTGAGGAACCTAGTACTAGAG
59
GATTC-AGGCTCAGAAGTCACAGGGAAACTGAGGCCTGGGTGAGGGTGT
56
ACATCCAGCTGGAAAAATCACCAGGGAGCTGGAGCTGATGGGATAAAAA
49
GGTACCAGGTTGAGCAGCTACAGGAGAGCTAGGACATGTGGGGATGTTT
69
TGTTCCAGGCTGAACAACTGTAGAGCATCAGGGGGAGGTGGAACTTTAA
63
GAAGTCAGGCTGAGCAGCTACAGGAGAGCTGCAGCTATTCGGTATGTGG
61
AGGTCCAGCCAGAGCAGCTACAGGGTAGCTGGGATAAATGGGGCTG-GA
71
GA-ACCA-AGCTAGAAGACACAGGGGAGCAGGTTCTAGTCTGCATAGGA
65
GTGGGGATCC
$S_{\gamma 3}$
GGTACCAAGTTGAGCAGCCACAGGAGAGCATAG----------
GGGACCTGGATAAGCCATTATGTGGGAGCTGGTGTAACTGGAAATATAG
67
GGAAATAGAATAACCAGCTACACAGGAGATCTAGAGGGAAGGAGCATAA
58
GGAGTCTGACCAAGCAACCATAGTGGGGCTGGGGAAGCTGAGAGTATGC
74
ACAGCCAAGCTGAGAAGTTAAAGGAGAACAGGGGTAGGTGAGGGTGTGA
63
AGTACCAGAA-TCTGAGCTACAGAGGAGCTGGGCAGG-TGGGAATATGG
77
AGGACCGGGTTGAGAATCCACAGAGAGCCACCAGGTGGCAGTTCCACAG
43

Figure 1


1
10
20
30
40
50
$S_{\gamma 1}$ 5′ GPPTCCAGGCTGAGCAGCTACAGGGGAGCTGGGGYAPPTGGGAPTPTPG 3′
$S_{\gamma 2b}$ 5′ GGGACCAGTCCTAGCAGCTPTGGGGGAGCTGGGGAATGGTPGGAPTPTGA 3′
$S_{\gamma 3}$ 5′ PGNACCAGPNTPAGCAPYYACAGGGGAGCTGGGGATPGGTGGGAGTATPP 3′


kb
a
b
c
d
e
f
g
h
i
kb
15
5.0
3.8
3.55
3.3
2.5
0.8
3.55
3.3
0.8



Igγ2b-23 Igγ2b-26

Gene, 11 (1980) 117–127

Nucleotide sequences of class-switch recombination region of the mouse immunoglobulin γ2b-chain gene

(Recombinant DNA; S region; μ chain; inverted-repeat-stem structure; bacterial insertion element IS*2*; mouse myeloma; plasmid pBR322; λ phage vector)

Naoki Takahashi *, Tohru Kataoka and Tasuku Honjo **

Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo 113 (Japan)

(Received April 7th, 1980)
(Accepted May 16th, 1980)

SUMMARY

We have cloned a rearranged γ2b-chain gene from mouse myeloma MOPC141. Nucleotide-sequence determination demonstrated that the rearranged γ2b-chain gene was formed by a recombination between the region flanking the 5′-ends of the embryonic μ-chain gene and that of the γ2b-chain gene. An inverted-repeat-stem structure can be formed at the junction of the embryonic μ- and γ2b-chain genes. Nucleotide sequences similar to the one located adjacent to the recombination site were found in other parts of the region flanking the 5′-end of the μ-chain-gene, as well as in the region flanking the 5′-end of the γ1-chain gene. The common sequence (GGTANNAAAGNAC) shares extensive homology with the sequence proximal to the insertion site of the bacterial insertion element IS*2*.

INTRODUCTION

Immunoglobulin proteins comprise two heavy (H) and two light (L) chains, each consisting of an amino-terminal variable (V) region and a carboxy-terminal constant (C) region. The H chains are classified into five classes (μ, δ, γ, α, and ε). During a course of differentiation a single lymphocyte appears to sequentially produce different classes of immunoglobulin proteins, which is called class-switch phenomenon. In a given lymphocyte, a single V region of the H-chain associates successively with different C regions, while the L-chain protein synthesized remains unchanged. The switch of the H chain follows the sequence from μ to γ or α.

In 1978 we proposed that the order of the H-chain genes is V_H, spacer, μ, γ3, γ1, γ2b, γ2a and α, and that in the process of class switch, a chromosomal segment which is located between recombining V_H and C_H genes is excised out from the chromosomes so as to bring C and V genes closer (Honjo and Kataoka, 1978). The class-switch phenomenon was explained

* On leave from Department of Agricultural Chemistry, Faculty of Agriculture, University of Tokyo, Tokyo 113 (Japan).

** To whom all correspondence should be sent at the Department of Genetics, Osaka University Medical School, Kita-ku, Osaka 530 (Japan).

Abbreviations: bp, base pairs; kb, kilobase pairs.

,y the successive recombination event of a V_H gene from μ to γ or α. The linear arrangement of C_H genes and their deletion in myeloma cells have recently oeen confirmed by several groups including our own (Coleclough et al., 1980; Cory and Adams, 1980; Cory et al., 1980; Rabbits et al., 1980; Yaoita and Honjo, 1980a, b).

We have recently cloned an 8.3 kb long *Eco*RI fragment containing a rearranged γ1-chain (IgH7) from a γ1-chain-producing myeloma MC101, and compared its structure with the embryonic γ1 (IgH2) and μ (IgH701) chain gene fragments, both cloned from newborn mouse DNA (Kataoka et al., 1980). Studies using restriction-site mapping and Southern hybridization of restriction fragments have ıdicated that the rearranged γ1-chain gene clone comprises a 6.3 kb long fragment of the embryonic γ1 gene and a newly introduced segment (2 kb) that was originally located at the flanking 5′-region of the embryonic μ-chain gene. The recombination site of the two segments seemed to be different from the putative J region, which is the recombination site, to bring a V_L gene to the close proximity of the C_L gene (Max et al., 1979; Sakano et al., 1979; Seidman et al., 1979). Based on these results we have extended the previous model (Honjo and Kataoka, 1978) and proposed that the class-switch recombination takes place between novel regions, called S regions, which are present in the flanking 5′-region of each C_H gene (Kataoka et al., 1980). We assume that putative J_H regions are clustered upstream of the S region of the μ-chain gene ($S\mu$), and that the recombination between V_H and J_H regions results in expression of the μ-chain gene. Then, the class switch from μ to γ1 is probably mediated by a second recombination event between the S_μ and $S_{\gamma 1}$ regions, without affecting the V-region sequence. Davis et al. (1980) analyzed the rearranged α-chain gene from α-chain-producing myeloma, M603 and have also shown that the complete α-chain gene is formed by two distinct types of DNA rearrangements, a V_H-J_H joining and a C_H switch.

To further examine this model we have cloned an *Eco*RI fragment (7.8 kb) containing the rearranged γ2b-chain gene from a γ2b-chain-producing myeloma MOPC141, and determined the nucleotide sequence around the recombination site of the rearranged γ2b-chain gene. This sequence was compared with those in the corresponding regions of the embryonic μ and γ2b-chain genes. The results unequivocally demonstrated that the rearranged γ2b-chain gene is formed by a recombination between a region flanking the 5′-end of the embryonic μ-chain gene and that of the γ2b-chain gene. The nucleotide sequences around the junction of the μ and γ2b-chain genes allow the formation of an inverted-repeat-stem structure, and share homology with that of the prokaryotic insertion element IS*2* (see Szybalski, 1977).

MATERIALS AND METHODS

(a) Materials

[γ-^{32}P]ATP (spec. act. >5500 Ci/mmol) and [α-^{32}P]dCTP (spec. act. >3000 Ci/mmol) were purchased from New England Nuclear (Boston, MA) or from Radiochemical Centre (Amersham, England). Sources of restriction endonucleases, polynucleotide kinase and DNA ligase have been described previously (Honjo et al., 1979; Obata et al., 1980).

High molecular weight DNA was prepared from MOPC141 myeloma tumor as described (Honjo et al., 1979). The myeloma MOPC141 was kindly provided by Dr. M. Potter of NIH and propagated subcutaneously in BALB/c mice. The myeloma tumors were harvested and stored at −80°C until used for DNA extraction.

(b) Preparation and screening of recombinant phage

Bacteriophage λgt*WES* · λB′ (Leder et al., 1977) was used as an EK2 vector and propagated in DP50-*supF*. Cloning experiments were carried out in a P2 facility. Ligation of λgt*WES* outer fragments and purified mouse DNA, and in vitro packaging of recombinant DNAs were carried out as described (Blattner et al., 1978; Kataoka et al., 1979). Five sets of about 50000 plaques were directly transferred from LB broth agar plate (25 X 25 cm) to nitrocellulose filters (Schleicher and Schuell, Keene, MA) as described (Benton and Davis, 1977). The filters were hybridized to ^{32}P-labeled IgH22 and autoradiographed. Recombinant phage DNA was prepared as described (Tiemeier et al., 1977). The insert DNA was isolated and recloned into pBR322 by ligation (Bolivar et al., 1977). Plasmid DNA was isolated as described (Clewell and Helinski, 1969).

(c) Southern blot hybridization

DNA was digested with appropriate restriction endonucleases, and electrophoresed in 0.8% agarose gels, which were blotted according to Southern (1975). The filter was hybridized to ^{32}P-labeled nick-translated DNA as described (Kataoka et al., 1979).

(d) Nucleotide sequence determination

Nucleotide sequence was determined according to the method of Maxam and Gilbert (1977) with slight modifications (Honjo et al., 1979).

TABLE I

Immunoglobulin gene clones used in this study

Cloned insert	Size (kb)	Structural gene	Source of DNA	Ref.
IgH22	6.6	γ2b	newborn mouse	Yamawaki-Kataoka et al. (1980)
IgH2201	7.8	γ2b	MOPC141	this paper
IgH701	13	μ	newborn mouse	Kataoka et al. (1980)
IgH2	6.6	γ1	newborn mouse	Honjo et al. (1979)
IgH7	8.3	γ1	MC101	Kataoka et al. (1980)

RESULTS AND DISCUSSION

(a) Cloning and characterization of the rearranged γ2b chain gene

Previously (Kataoka et al., 1980; Yaoita and Honjo, 1980a,b) we examined restriction fragments of the γ2b-chain gene in *Eco*RI-digested DNAs of various myelomas as well as newborn mice using the nick-translated γ2b-chain gene fragment (IgH22) cloned from *Eco*RI-digests of newborn mouse DNA (Kataoka et al., 1979; Yamawaki-Kataoka et al., 1980) as a hybridization probe. DNAs of newborn mice and myelomas producing the μ, γ3 and γ1 chains contained a single γ2b-chain gene fragment (6.6 kb) which is indistinguishable from IgH22. When digested with *Eco*RI, DNA of a γ2b-chain-producing myeloma MOPC141 produced an extra γ2b-gene band of 7.8 kb in addition to the 6.6 kb band (Kataoka et al., 1980; Yaoita and Honjo, 1980a,b).

The *Eco*RI fragment (7.8 kb) of the rearranged γ2b-chain gene of MOPC141 DNA was partially purified by a preparative agarose gel (Polsky et al., 1978) and cloned using λgt*WES* as an EK2 vector (Leder et al., 1977). Two independent positive clones were obtained after screening 2.5×10^5 hybrid phages. These two clones were indistinguishable by restriction site analysis and designated as λgt*WES*::IgH2201, the insert of which is called IgH2201. Immunoglobulin gene clones used in the present study are summarized in Table I.

IgH2201 and IgH22 DNAs were digested with various restriction enzymes and the resultant DNA fragments were compared. The restriction maps of IgH2201 (7.8 kb) and IgH22 (6.6 kb) (Yamawaki-Kataoka et al., 1980) are shown in Fig. 1. The 3′-side regions (6.5 kb) of IgH2201 and IgH22 were indistinguishable from each other by restriction cleavage sites. The remaining 5′-side regions of IgH2201 and IgH22 were different from each other with respect to locations of various restriction sites. We have tested whether or not the rearranged γ2b-gene fragment (IgH2201) contains a portion of the embryonic μ-chain gene fragment (IgH701) cloned from the *Eco*RI digests of newborn mouse DNA since we knov that the 5′-portion of the rearranged γ1-chain gene fragment is derived from the 5′-portion of IgH701 (Kataoka et al., 1980). Comparison of the restriction cleavage maps of IgH2201 and IgH701 indicates that the 5′-terminal 1.3 kb region of IgH2201 is identical to that of IgH701 (Fig. 1).

Comparison of the restriction cleavage maps of IgH701, IgH22 and IgH2201 indicates that IgH2201 (7.8 kb) consists of the 5′-terminal portion (1.3 kb) of IgH701 and the 3′-terminal portion (6.5 kb) of IgH22. Restriction DNA fragments of IgH701 and IgH22 hybridized with nick-translated IgH2201, an observation consistent with the above conclusion (Fig. 2). λgt*WES*::IgH701 DNA, when digested with *Hin*dIII/*Eco*RI, *Hha*I/*Eco*RI and *Xba*I/*Eco*RI, produced 0.7 (doublet), 1.4 and 6.0 kb fragments, respectively, which hybridized with ^{32}P-labeled IgH2201 (Fig. 2). All of these fragments contain, at least, a part of the 5′-terminal portion (1.3 kb) of IgH701 (Fig. 1). A

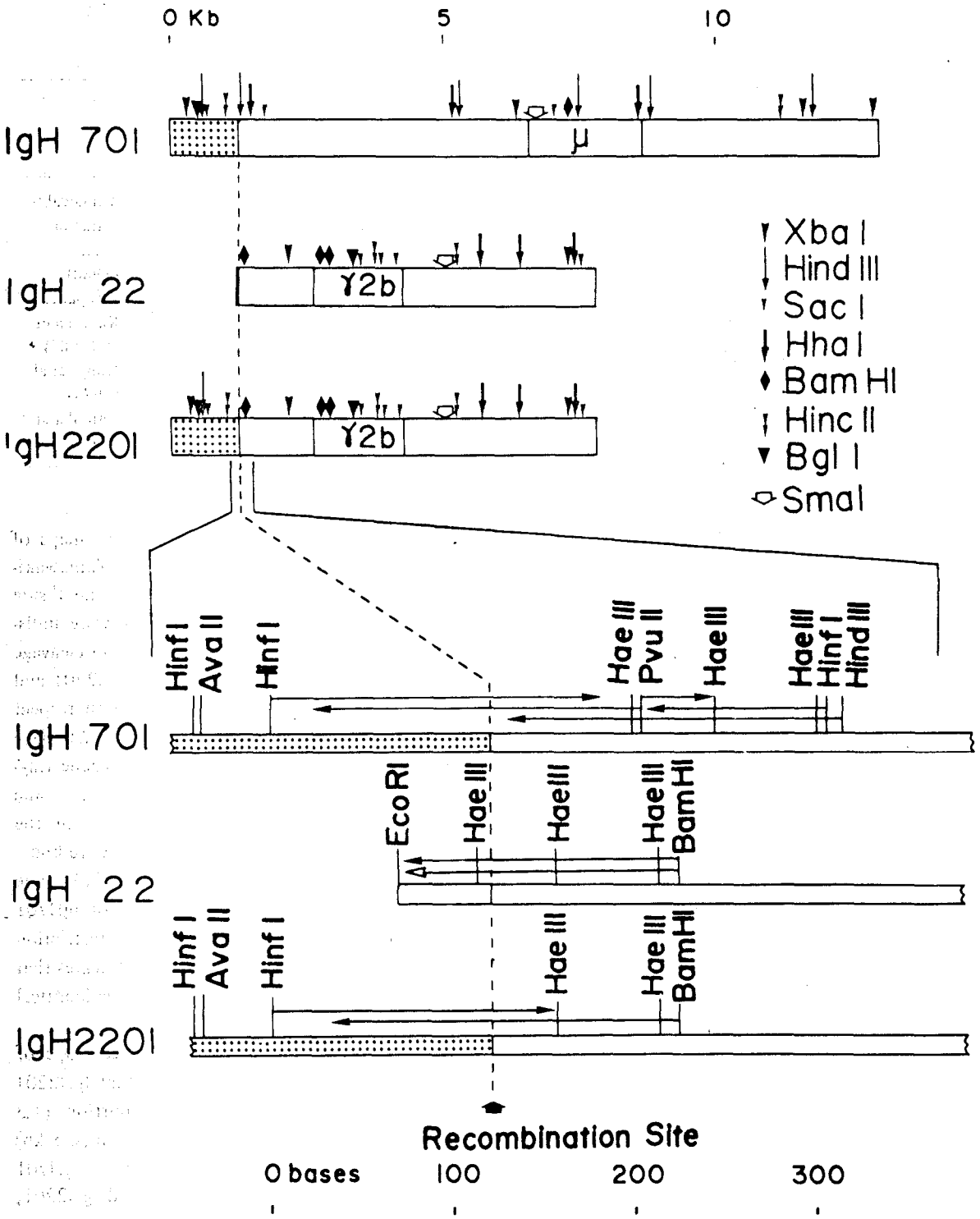


Fig. 1. Comparison of restriction enzyme cleavage maps of IgH701, IgH22 and IgH2201 and sequencing strategy of the recombination sites of IgH701, IgH22 and IgH2201. Cloned DNAs (*Eco*RI fragments) are displayed with the direction of transcription from left to right. The restriction cleavage sites were determined by combined cleavages with two or more enzymes. Restriction enzyme cleavage sites of IgH701 and IgH22 were determined previously (Kataoka et al., 1980; Yamawaki-Kataoka, 1980). The dotted line shows the recombination site. The dotted regions represent the segment originally present at the 5'-terminal portion of IgH701. The restriction sites used for sequencing are indicated in the enlarged map shown below. The ranges and directions of the sequences read are indicated by horizontal arrows. An arrow with a white head indicates that the DNA segment was labeled at its 3'-end by the reverse transcriptase.

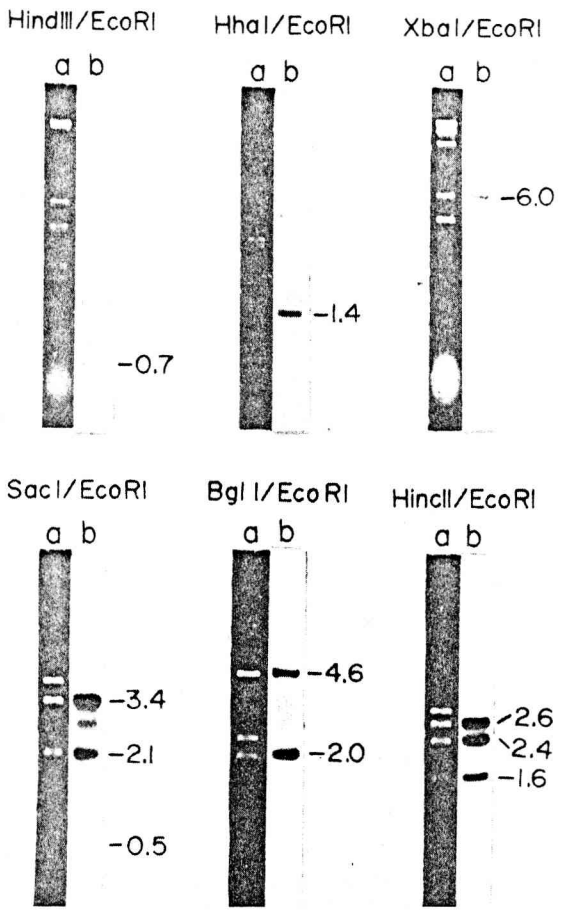


Fig. 2. Southern hybridization of λgt*WES*::IgH701 and pBR322::IgH22 restriction DNA fragments with ^{32}P-labeled IgH2201 DNA. Approx. 0.5 μg each of λgt*WES*::IgH701 and pBR322::IgH22 DNAs was cleaved with restriction enzymes indicated and electrophoresed in 0.8% agarose gels. The Southern blots of the gels were hybridized with ^{32}P-labeled IgH2201 DNA. The following pairs of lanes are ethidium bromide stains (lane a) and autoradiograms (lane b) of Southern blots for cleavage of λgt*WES*::IgH701 DNA by *Hin*dIII/*Eco*RI, *Hha*I/*Eco*RI and *Xba*I/*Eco*RI and of pBR322::IgH22 DNA by *Sac*I/*Eco*RI, *Bgl*I/*Eco*RI and *Hin*cII/*Eco*RI. The sizes of fragments are shown in kb.

small *Xba*I/*Eco*RI fragment (0.3 kb), which is at the 5′-end of IgH2201, hybridized weakly with [^{32}P]-IgH2201 and the band was visible only after very long exposure. pBR322 :: IgH22 DNA, when digested with *Sac*I/*Eco*RI, produced three fragments of 3.4, 2.1 and 0.5 kb, which hybridized with ^{32}P-labeled IgH2201 (Fig. 2). These fragments comprise a major portion of IgH22 (Fig. 1). Small *Sac*I/*Eco*RI fragments (0.3 kb doublet) hybridized weakly with the probe and the bands were visible only after long exposure. Similarly, *Bgl*I/*Eco*RI and *Hin*cII/*Eco*RI digestions of pBR322 :: IgH22 produced two fragments (4.6 and 2.0 kb) and three fragments (2.6, 2.4 and 1.6 kb), respectively, which hybridized with ^{32}P-labeled IgH2201 (Fig. 2). These fragments represent almost all of the IgH22 DNA (Fig. 1). These results lead us to conclude that the DNA segment present at the 5′ proximity of the rearranged γ2b-chain gene in the γ2b-chain-producing myeloma MOPC141 has been derived from the flanking 5′-region of the μ-chain gene in newborn mouse DNA.

(b) Nucleotide sequence around recombination sites of embryonic and rearranged γ2b chain genes

The detailed restriction maps and the sequencing strategy around the recombination sites of IgH701, IgH22 and IgH2201 are also shown in Fig. 1. IgH22 and IgH2201 share the common restriction enzyme cleavage pattern to the right of the *Hae*III sites that lie 0.09 and 1.26 kb downstream of the 5′-ends of IgH22 and IgH2201, respectively. However, a *Hae*III site 0.05 kb away from the 5′-end of IgH22 is absent from IgH2201 at the corresponding position. Upstream from this *Hae*III site, the cleavage pattern of IgH2201 is quite different from that of IgH22, but is identical to that of IgH701.

The nucleotide sequences around the recombination sites of IgH701, IgH22 and IgH2201 are shown in Fig. 3. Comparison of these sequences locates th recombination site in the rearranged γ2b-chain gene fragment at base 100 or 101 before the *Bam*HI site. The nucleotide sequences of IgH22 and IgH2201 are identical downstream of the recombination site, whereas IgH2201 shares the common nucleotide sequence with IgH701 upstream of the recombination site. The nucleotide sequence around the recombination site does not contain a sequence similar to the putative V_H or J_H sequences, which are deduced from the amino acid sequence data (Fougereau et al., 1976; Barstad et al., 1974; Adetsugbo, 1978; Kabat et al., 1978; Rao et al., 1979). The results unequivocally demonstrate that the rearranged γ2b-chain gene is formed by a recombination between a region flanking the 5′-end of the embryonic μ-chain gene and a region flanking the 5′-end of the embryonic γ2b-chain gene. The recombination site is different from the putative J_H region, in agreement with the previous studies on the rearranged γ1-chain gene (Kataoka et al., 1980).

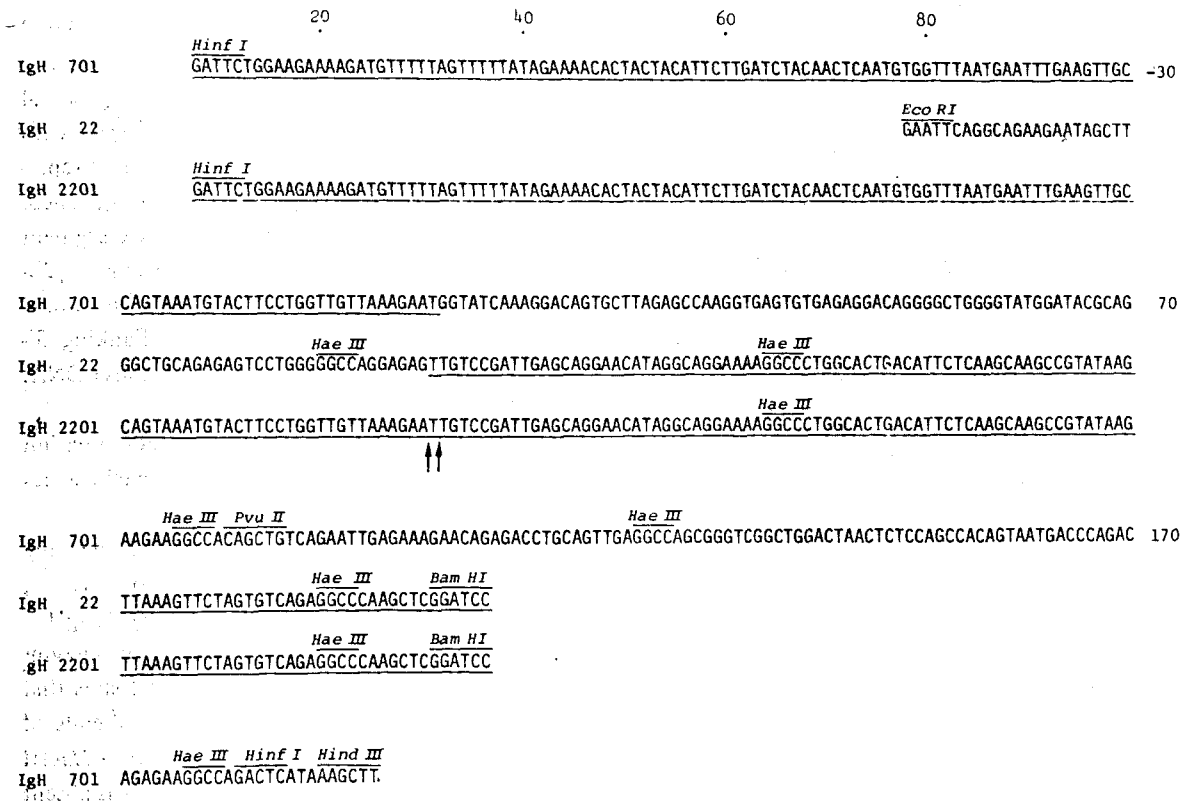


Fig. 3. Nucleotide sequences around recombination sites of IgH701, IgH22 and IgH2201. DNA fragments to be sequenced were isolated as follows. (1) IgH701: *Hin*dIII-E fragment (0.7 kb) of IgH701 was subcloned into pBR322 plasmid. The *Hin*dIII-E fragment was isolated by 3% polyacrylamide gel electrophoresis after digestion of the plasmid DNA with *Hin*dIII, digested with either *Pvu*II or *Hin*fI and used for the substrates of phosphorylation. (2) IgH22: The 0.45 kb *Bam*HI fragment containing the 0.3 kb portion of pBR322 and the 5′-terminal 0.15 kb *Bam*HI-*Eco*RI portion of IgH22 was isolated by 4% polyacrylamide gel electrophoresis after digestion of pBR322::IgH22 with *Bam*HI and its 5′-ends were labeled by phosphorylation. The 5′-terminal 0.15 kb *Bam*HI-*Eco*RI fragment was also isolated and its 3′-end was labeled by [α-^{32}P]dCTP reverse transcriptase. (3) IgH2201: The 0.6 kb *Hin*dIII-*Bam*HI fragment was isolated by 3% polyacrylamide gel electrophoresis after digestion of IgH2201 DNA with *Hin*dIII and *Bam*HI. The 0.6 kb fragment and its *Hin*fI digests were used for the substrates of phosphorylation. Nucleotide sequences were displayed from left to right with the direction of transcription of the structural sequence. Nucleotide sequences are underlined where pairs of clones have the identical nucleotide sequences. Restriction-enzyme recognition sequences have a line drawn above hem. Arrows indicate the recombination sites.

Such recombination may be responsible for class-switch phenomenon since rearrangement of H chain genes takes place specifically at the flanking 5′-regions of the expressed H chain genes (Kataoka et al., 1980; Yaoita and Honjo, 1980a,b). We postulate that IgH2201 is the 3′-fragment of the completed γ2b-chain gene that is linked to the V_H gene and expressed in MOPC141 myeloma.

Possible inverted-repeat-stem structure around recombination site

An inverted-repeat-stem structure could be formed between the nucleotide sequence present on the 5′-side of the recombination site of the embryonic γ2b-chain gene and that present on the 3′-side of the recombination site of the embryonic μ-chain gene as shown in Fig. 4. The specific nucleotides that were joined by the recombination event are shown by arrows in Fig. 4. These nucleotides lie near the base of the hypothetical stem structure, and the stem would be eliminated by the recombination event. According to the proposed order of C_H genes (Honjo and Kataoka, 1978) the looped-out region between IgH701 and IgH22 contains the structural genes of the μ, γ3, and γ1 chains. We have recently shown that DNA of MOPC141 myeloma is devoid of the γ1-chain gene (Yaoita and Honjo, 1980a,b).

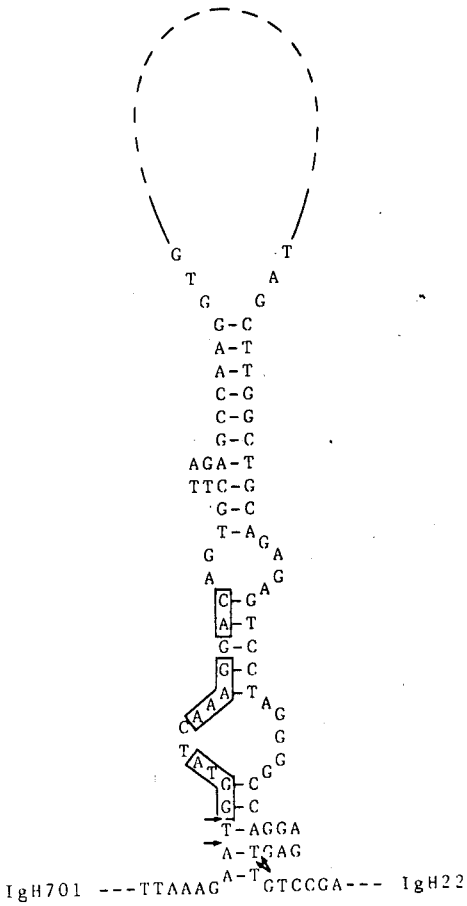


Fig. 4. Looping out model for recombination of IgH701 and IgH22. The nucleotide sequence common to many recombination sites (see Fig. 5) is framed. Arrows indicate the recombination sites.

A similar stem structure was assumed to form between the 3′-end of an embryonic V_L sequence and the 5′-end of a J_L sequence. The nucleotide sequences in the stem structures, however, are different between the μ-γ2b and the V_L-J_L joints (Max et al., 1979; Sakano et al., 1979).

The nucleotide sequence present on the 3′-side of the μ-γ2b recombination site, representing a portion of the μ-chain gene fragment (IgH701), shares extensive homology with the sequence adjacent to the recombination site of the γ1-chain gene fragment published previously (Kataoka et al., 1980) (Fig. 5, A and B). We also found a third similar sequence at about 0.33 kb downstream of the μ-γ2b recombination site of IgH701 (Fig. 5, C). This sequence is present at the 3′-side adjacent to the IgH701-IgH703 recombination site (see below) of IgH701 (Kataoka et al., 1980; Y. Yamawaki-Kataoka and T. Honjo, unpublished data). We have then looked for yet other similar sequences which may be present elsewhere in IgH701. To do this we determined the nucleotide sequence of the 5′-segment of the rearranged γ1-chain gene fragment (IgH7) cloned from *Eco*RI digest of MC101 DNA (Table I). This segment has been shown to be derived from the 5′ portion of the embryonic μ-chain gene clone (IgH701) (Kataoka et al., 1980). The strategy of nucleotide sequence determination and the nucleotide sequence of IgH7 are shown in Fig. 6. This sequence is essentially identical to that of the corresponding region (about 600 bp) of IgH701, whose sequence was determined independently. We found another similar sequence 608 bases upstream of the μ-γ2b recombination site (Fig. 5, D). These similar sequences are listed in Fig. 5.

Furthermore, it is interesting to point out that these sequences share extensive homology with that of the *gal* promoter region where the bacterial insertion element IS*2* is inserted (Fig. 5, IS*2*) (Musso and Rosenberg, 1977). The common sequence is GGTANNAAAGNAC. Note that this sequence is present at the neck of the stem structure shown in Fig. 4. It is likely that such a sequence comprises, at least in part, the recognition signal for the immunoglobulin gene recombination.

(d) S regions

We proposed that the class-switch recombination takes place by the mutual recognition of two S regions, each of which is located proximal to different C_H genes (Kataoka et al., 1980). Since the S region does not code for any polypeptide, it is hard to pinpoint the exact nucleotide sequences of the S regions. We find nucleotide sequences, which share GGTANNAAAGNAC in common, are present close to the S-mediated recombination sites (A, C and B in Fig. 5). At least three sets of sequences of this type seem to be located in the flanking 5′-region of the μ-chain gene (A, C and D in Fig. 5). However, the nucleotide sequences around the recombination sites of the rearranged γ1-chain gene fragment (IgH7) derived from MC101 myeloma and of the rearranged γ2b-chain gene fragment (IgH2201) derived from MOPC141 myeloma are different from each other.

We have found tandem repetitive sequences of (GAGCT)n, (GAGAG)n, and (GGCTG)n in the

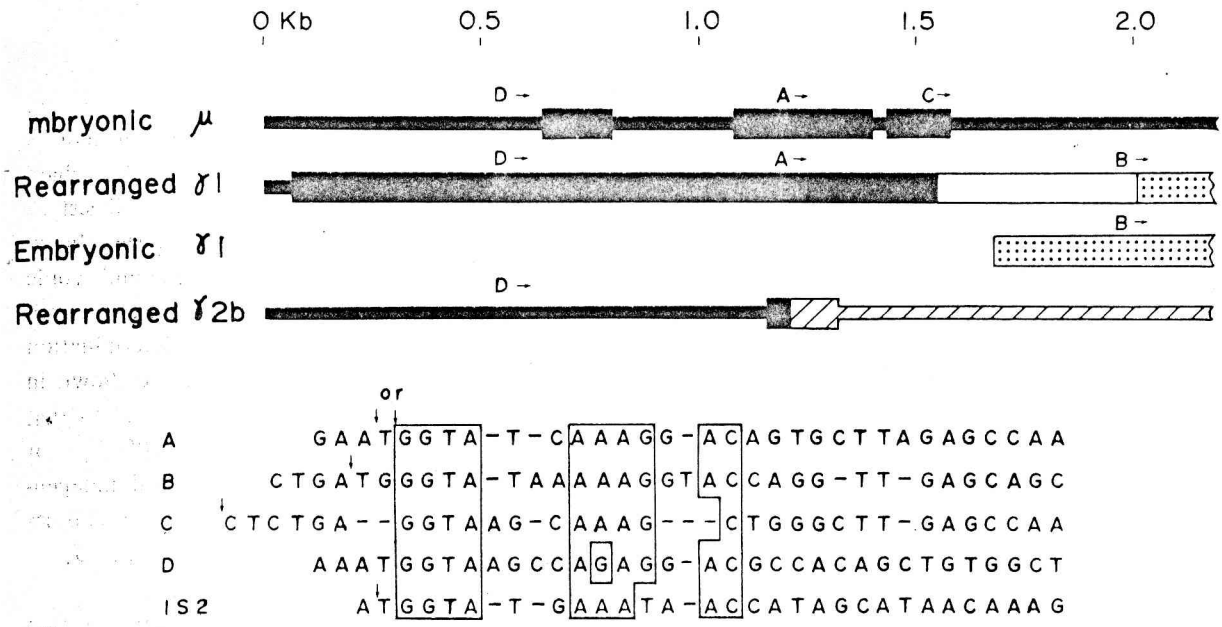


Fig. 5. Nucleotide sequences homologous to that of the 3′-region adjacent to the recombination site of IgH701. Horizontal arrows indicate the location and direction of the homologous sequences. Coding strand of the IS2 insertion site in the *gal* promoter region is displayed with the IS2 on the left (Musso and Rosenberg, 1977). Vertical arrows indicate the site of recombination. The common sequences are framed. Closed, dotted, hatched and open rectangles indicate the segments of DNA originally present in the embryonic type of the μ-chain gene (IgH701), the γ1-chain gene (IgH2), the γ2b-chain gene (IgH22) and an unknown gene (IgH703), respectively. Wider rectangles indicate the sequenced regions.

400-bp-long region before the recombination site in IgH7 (Kataoka et al., 1980). Although there are no obvious repetitive sequences in the immediate proximity of the recombination site in IgH2201, repetitive sequences are found in the region 2–4 kb downstream from the 5′-end of the embryonic μ-chain gene clone (T. Kawakami and T. Honjo, unpublished data). We have recently cloned another rearranged γ2b-chain gene fragment from MPC11 myeloma and found that this DNA carries long tandem repetitive sequences close to the recombination site (T. Kataoka and T. Honjo, unpublished data). Although the nucleotide sequences of these repetitive sequences are different among the γ1-, γ2b- and μ-chain genes, the presence of such tandem repetition may help to enhance recombination events.

Since the class-switch-recombination site(s) is located in the intervening sequence between a V_H gene and a C_H gene, the S-S recombination event does not have to be highly specific to the nucleotides that are joined together. Instead, the S-mediated recombination is expected to be efficient, because in the small B lymphocyte the H chain class is switched from μ to γ or to α within a short period of time after encountering an antigen. In this respect, the S-S recombination is in sharp contrast to the V-J recombination, which requires extremely specific excision and ligation. From these points of view, it is possible to imagine that a combined structure of tandem repetitive sequences and relatively short common sequences may serve as the S region. However, it is premature to pinpoint the S region at this stage.

(e) Comparison between rearranged γ1- and γ2b-chain genes

We have reported that the rearranged γ1-chain gene fragment (IgH7) cloned from MC101 myeloma DNA contains, in addition to the embryonic γ1-chain gene fragment, DNA segments derived from not only the μ chain gene clone (IgH701) but also a second embryonic gene clone IgH703 (Kataoka et al., 1980). The IgH703 may represent a part of or a region proximal to, a H chain gene (possibly δ-, γ3- or ε-chain genes) present between the μ- and γ1-chain genes. In MC101 myeloma, therefore, the class

Fig. 6. Sequence of IgH7 and strategy of nucleotide-sequence determination. IgH7 DNA was recloned into a plasmid pBR322 (pIgH7). Restriction cleavage sites were mapped by combined cleavage of the plasmid DNA with various restriction enzymes, or partial digestion of DNA fragments labeled at one of their 5′-termini with T4 polynucleotide kinase and [γ-^{32}P]ATP (Smith and Birnstiel, 1976). The 2.6 kb *Bam*HI fragment containing the 5′-terminal 2.2 kb DNA segment of IgH7 and 0.38 kb segment derived from pBR322 and the 1.3 kb *Sac*I-*Kpn*I fragment were isolated by 3% polyacrylamide gel electrophoresis. Each fragment was digested with the second restriction enzyme and used as substrate for phosphorylation. Nucleotide sequence was displayed from left to right with the direction of transcription of the structural sequence. Position numbers are matched to those in Fig. 3. The nucleotide sequences similar to the recombination site sequence are framed (see Fig. 5, A and D). Strategy for nucleotide sequence determination is shown below the sequence. Arrows indicate the ranges and directions of the sequence read.

switch may have proceeded from μ to γ1 by two successive S-S recombinations. On the other hand, we saw that the rearranged γ2b-chain gene clone (IgH2201) consists of the embryonic γ2b-chain gene and the segment of the embryonic μ-chain gene clone. In MOPC141 myeloma cells, the class switch might have proceeded directly from μ to γ2b, though we cannot exclude the possibility that the original rearranged γ2b-chain gene had contained a DNA segment derived from the proximal region of another H-chain gene, which was later deleted.

ACKNOWLEDGEMENTS

We thank Dr. Y. Mano for his continuous encouragement. This investigation was supported in part by grants from the Ministry of Education, Science and Culture of Japan.

REFERENCES

Adetsugbo, K.: Evolution of immunoglobulin subclasses: primary structure of a murine myeloma γ1 chain. J. Biol. Chem. 253 (1978) 6068–6075.

Barstad, P., Farnsworth, V., Weigert, M., Cohn, M. and Hood, L.: Mouse immunoglobulin heavy chains are coded by multiple germ-line variable region genes. Proc. Natl. Acad. Sci. USA 71 (1974) 4096–4100.

Benton, W.D. and Davis, R.W.: Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196 (1977) 180–182.

Blattner, F.R., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Richard, J.E., Slighton, J.L., Tucker, P.W. and Smithies, O.: Cloning human fetal γ-globin and mouse α-type globin DNA: preparation and screening of shotgun collections. Science 202 (1978) 1279–1283.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2 (1977) 95–113.

Clewell, D. and Helinski, O.: Supercoiled circular DNA-protein complex in *Escherichia coli:* purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA 62 (1969) 1159–1166.

Coleclough, C., Cooper, C. and Perry, R.P.: Rearrangement of immunoglobulin heavy chain genes during B-lymphocyte development as revealed by studies of mouse plasma cells. Proc. Natl. Acad. Sci. USA 77 (1980) 1422–1426.

Cory, S. and Adams, J.M.: Deletions are associated with somatic rearrangement of immunoglobulin heavy chain genes. Cell 19 (1980) 37–51.

Cory, S., Jackson, J. and Adams, J.M.: Deletions within the constant region locus can account for switches in immunoglobulin heavy chain expression. Nature (1980) in press.

Davis, M.M., Calame, K., Early, P.W., Livant, D.L., Joho, R., Weissman, I.L. and Hood, L.: An immunoglobulin heavy-chain gene is formed by at least two recombinational events. Nature 283 (1980) 733–739.

Fougereau, M., Bourgois, A., de Preval, C., Rocca-Serra, J. and Schiff, C.: The complete sequence of the murine monoclonal immunoglobulin MOPC173 (IgG2a): genetic implications. Ann. Immunol. (Paris) 127C (1976) 607–631.

Honjo, T. and Kataoka, T.: Organization of immunoglobulin heavy chain genes and allelic deletion model. Proc. Natl. Acad. Sci. USA 75 (1978) 2140–2144.

Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N. and Mano, Y.: Cloning and complete nucleotide sequence of mouse immunoglobulin γ1 chain gene. Cell 18 (1979) 559–568.

Kabat, E.A., Wu, T.T. and Bilofsky, H.: Variable region genes for the immunoglobulin framework are assembled from small segments of DNA – a hypothesis. Proc. Natl. Acad. Sci. USA 75 (1978) 2429–2433.

Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. and Honjo, T.: Cloning immunoglobulin γ2b chain gene of mouse: characterization and partial sequence determination. Proc. Natl. Acad. Sci. USA 76 (1979) 4240–4244.

Kataoka, T., Kawakami, T., Takahashi, N. and Honjo, T.: Rearrangement of immunoglobulin γ1-chain gene and mechanism for heavy-chain class switch. Proc. Natl. Acad. Sci. USA 77 (1980) 919–923.

Leder, P., Tiemeier, D. and Enquist, L.: EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the λgt*WES* system. Science 196 (1977) 175–177.

Max, E.E., Seidman, J.G. and Leder, P.: Sequences of five potential recombination sites encoded close to an immunoglobulin κ constant region gene: Proc. Natl. Acad. Sci. USA 76 (1979) 3450–3454.

Maxam, A.M. and Gilbert, W., A new method for sequencing DNA: Proc. Natl. Acad. Sci. USA 74 (1977) 560–564.

Musso, R.E. and Rosenberg, M.: Nucleotide sequences at two sites for IS2 DNA insertion, in Bukhari, A.I., Shapiro, J.A. and Adhya, S.L. (Eds.), DNA Insertion Elements, Plasmids, and Episomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 597–600.

Obata, M., Yamawaki-Kataoka, Y., Takahashi, N., Kataoka, T., Shimizu, A., Mano, Y., Seidman, J.G., Peterlin, B.M., Leder, P. and Honjo, T.: Immunoglobulin γ1 heavy chain gene: structural gene sequences cloned in a bacterial plasmid. Gene 9 (1980) 87–97.

Polsky, F., Edgell, M.H., Seidman, J.G. and Leder, P.: High capacity gel preparative electrophoresis for purification of

fragments of genomic DNA, Anal. Biochem. 87 (1978) 397–410.

Rabbitts, T.H., Forster, A., Dunnick, W. and Bentley, D.L.: The role of gene deletion in the immunoglobulin heavy chain switch. Nature 283 (1980) 351–356.

Rao, D.N., Rudikoff, S., Krutzsch, H. and Potter, M.: Structural evidence for independent joining region gene in immunoglobulin heavy chain from anti-galactan myeloma proteins and its potential role in generating diversity in complementarity-determining regions. Proc. Natl. Acad. Sci. USA 76 (1979) 2890–2894.

Sakano, H., Hüppi, K., Heinrich, G. and Tonegawa, S.: Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 280 (1979) 288–294.

Seidman, J.G., Max, E.E. and Leder, P.: A κ-immunoglobulin gene is formed by site-specific recombination without further somatic mutation. Nature 280 (1979) 370–375.

Smith, H.O. and Birnstiel, M.L.: A simple method for DNA restriction site mapping. Nucl. Acids Res. 3 (1976) 2387–2398.

Southern, E.M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis: J. Mol. Biol. 98 (1975) 503–517.

Szybalski, W.: IS elements in *Escherichia coli*, plasmids and bacteriophages, in Bukhari, A.I., Shapiro, J.A. and Adhya, S.L. (Eds.), DNA Insertion Elements, Plasmids, and Episomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 583–590.

Tiemeier, D.C., Tilghman, S.M. and Leder, P.: Purification and cloning of a mouse ribosomal gene fragment in coliphage lambda: Gene 2 (1977) 173–191.

Yamawaki-Kakaoka, Y., Kataoka, T., Takahashi, N., Obata, M. and Honjo, T.: Complete nucleotide sequence of immunoglobulin γ2b chain gene cloned from newborn mouse DNA. Nature 283 (1980) 786–789.

Yaoita, Y. and Honjo, T.: Deletion of immunoglobulin heavy chain genes accompanies the class switch rearrangement. Biomed. Res. 1 (1980a) 164–175.

Yaoita, Y. and Honjo, T.: Deletion of immunoglobulin heavy chain genes from expressed allelic chromosome. Nature (1980b) in press.

Communicated by K. Matsubara.

Reprinted from Nature, Vol. 289, No. 5794, pp. 149-153, January 15 1981

Ordering of mouse immunoglobulin heavy chain genes by molecular cloning

Akira Shimizu, Naoki Takahashi, Yuriko Yamawaki-Kataoka, Yasuyoshi Nishida, Tohru Kataoka & Tasuku Honjo

Department of Genetics, Osaka University Medical School, Kita-ku, Osaka 530, Japan

We have determined the order of the mouse immunoglobulin γ1, γ2b, γ2a and ε genes by molecular cloning of overlapping chromosomal segments. The results clearly demonstrate that the order is 5′–γ1–(21 kilobases)–γ2b–(15 kilobases)–γ2a–(15 kilobases)–ε–3′. There seem to be no J regions at the 5′ side of each constant region gene so far obtained except for the μ gene and these constant region genes seem to have repetitive sequences characteristic of switch (S) regions at their 5′ side.

IMMUNOGLOBULIN heavy (H) chain genes comprise a family of variable region (V) genes and several constant region (C) genes which are classified, in mouse, into five major classes: μ, γ, α, δ and ε. Genetic studies using allotype markers have clearly demonstrated that the V_H and C_H genes of a single chromosome are coordinately expressed—this is called *cis* expression[1-3]. During differentiation, a given lymphocyte seems to be able to associate a single V region sequentially with two or more different classes of H chain C region[4-11]. In this H chain class switch, the direction of the switch is always from μ to γ or α.

Using cDNA–mRNA hybridization kinetics, we have shown that specific C_H genes are deleted in mouse myelomas depending on which C_H genes are expressed[12], and have proposed that a chromosomal segment which intervenes recombining V_H and C_H genes is deleted to bring the V_H and C_H genes close together. Based on the deletion profiles of C_H genes in various myeloma tumours producing different C_H genes, a linear arrangement of C_H genes has been proposed as 5′–μ–γ3–γ1–γ2b–γ2a–α–3′. Subsequently, more direct evidence has been presented that deletion of the intervening DNA segment accompanies the H chain class switch[13-18]. Such studies using Southern blot[19] hybridization experiments also support the proposed order of C_H genes.

Comparative studies of rearranged and germ-line H chain genes have led to proposals[20-26] for a molecular mechanism of H chain class switch. According to this model, a complete H chain gene is formed by at least two types of the recombination event. The first type of recombination takes place between a given V_H, a J_H and a D gene segment, completing a V region gene[23-25]. The second type of recombination is required to switch a C_H gene and occurs between S_μ and S_γ (or S_α) regions which are located at the 5′ flanking regions of respective C_H genes. This model postulates two important structural features in the H chain gene organization: (1) the presence of only one set of J_H region genes in the 5′ side to the μ gene and (2) the presence of the S region before each C_H gene.

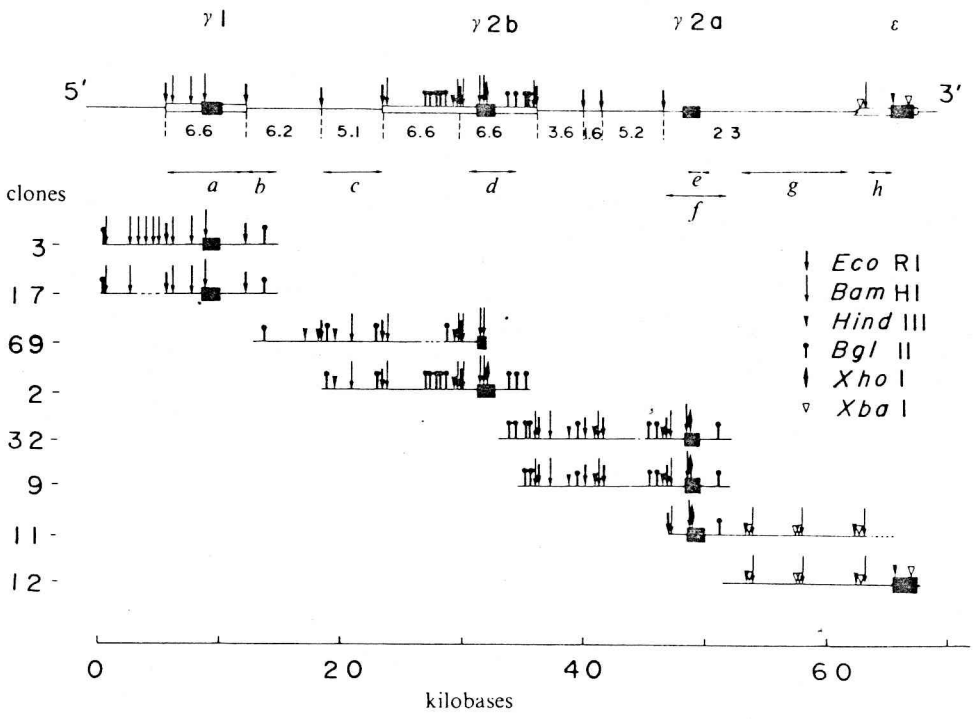


Fig. 1 Restriction endonuclease cleavage maps of overlapping cloned fragments between γ1 and ε genes. We used a recombinant phage library[42] for cloning DNA segments that contain immunoglobulin genes or their flanking regions, except for clone 11. The Charon 4A phage[43] library of BALB/c mouse embryo DNA (provided by Dr P. Leder) contained partial *Hae*III digests of embryonic mouse DNA with average length 18 kilobases. These fragments were ligated with *Eco*RI linkers and incorporated into phage DNA. The 20-kilobase insert of clone 11 was partially purified from *Eco*RI digests of RPC5 myeloma DNA by RPC5 column chromatography and agarose gel electrophoresis as described previously[27]. The partially purified DNA fragments were ligated with the outer fragments of Charon 4A phage DNA. Screening was done on 230×230 mm tryptone broth plates containing 5–20×10⁴ plaques by an *in situ* hybridization method[44]. Appropriate restriction endonuclease fragments of immunoglobulin genes were labelled with ^{32}P by nick translation[45] to a specific activity of 30–200 c.p.m. per pg and used as hybridization probes. Nitrocellulose filters were washed as described elsewhere[17]. A plaque containing an immunoglobulin gene or its flanking region was picked up and the phages were grown in liquid cultures to prepare DNA as described[46]. At the top is a schematic representation of the chromosomal segment containing the γ1, γ2b, γ2a and ε genes. Open bars indicate DNA segments cloned previously: the γ1 gene (IgH2)[27], the γ2b gene (IgH22, Igγ2b-26)[28,30,33] and the ε gene (Igε − 1)[34]. Structural genes are shown as solid boxes with direction of transcription from left to right, determined by nucleotide sequence determination. Numbers under the top line indicate sizes of *Eco*RI fragments in the germ-line gene. Horizontal arrows indicate fragments used as probes for screening or Southern hybridization. The insert of each clone is shown by a horizontal line with restriction sites. Clones were aligned by restriction endonuclease cleavage mapping and Southern hybridization analysis (see text and Figs 2, 4). Restriction digests were examined by electrophoresis in 0.5–0.7% agarose gels. It is possible that very small fragments (for example, 100 base pairs) may have been missed. *Bgl*II and *Xho*I sites at the 3′ side of the γ2a gene and *Xba*I sites at the 5′ side of the γ2a gene were not determined. Broken lines in clones 11, 17, 32 and 69 indicate deletions.

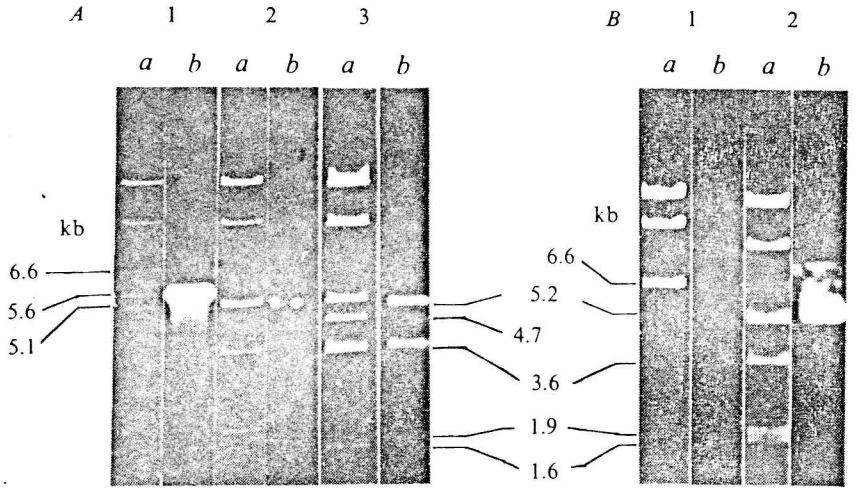


Fig. 2 Southern hybridization of isolated clones linking between the γ2b and γ2a genes. DNAs from isolated clones (0.1–0.3 μg each) were digested with *Eco*RI and electrophoresed in 0.5% agarose gels. The fragments were transferred to nitrocellulose filters[19] and hybridized with ^{32}P-labelled probes as described elsewhere[17]. Ethidium bromide stains (*a*) and autoradiograms (*b*) of Southern blots are shown. In *A*, 4.0-kilobase (kb) *Xba*I–*Hha*I fragment of γ2b gene (Fig. 1*d*) was used as a probe. Lanes 1, 2 and 3 contain DNAs of clones 2, 9 and 32, respectively. In *B*, γ2a cDNA clone (pG2a–14) (probe *e* in Fig. 1) was used as a probe. Lanes 1 and 2 contain DNAs of the 6.6-kilobase γ2b gene clone (IgH22)[28,30] and clone 9, respectively.

To determine directly the order of C_H genes and test the proposed C_H gene organization, we and others cloned mouse immunoglobulin genes, and have determined the complete nucleotide sequences of the γ1, γ2b and μ genes[27–31]. Liu *et al.*[32] isolated a clone which contains both μ and δ genes, and demonstrated that the μ gene is located ~4.5 kilobases 5′ to the δ gene. The linkage of 5′–γ1–γ2b–3′ was shown by cloning overlapping chromosomal segments[20]. We have recently cloned the γ3 and ε genes[33,34], and using these cloned DNA segments as probe, cloned overlapping chromosomal segments from embryonic mouse DNA. Characterization of these clones unequivocally demonstrates that the mouse immunoglobulin γ chain genes are aligned as proposed[12] (that is, 5′–γ1–γ2b–γ2a–ε–3′) on a chromosome. The linkage of 5′–ε–α–3′ is described elsewhere[34].

The γ2b gene is located 5′ to the γ2a gene

When we screened a Charon 4A library containing partial *Hae*III digests of embryonic mouse DNA with the 4-kilobase *Xba*I–*Hha*I fragment (probe *d* in Fig. 1) of the γ2b gene clone[28,30] as a probe, we isolated three clones which were designated Ch · M · Igγ2b–2 (clone 2), Ch · M · Igγ2a–9 (clone 9) and Ch · M · Igγ2a–32 (clone 32). These were characterized by restriction enzyme cleavage mapping and Southern blot hybridization (Figs 1, 2). Clone 2 contains an *Eco*RI fragment (5.6 kilobases) that hybridizes with probe *d* (Fig. 2*A*, lane 1) and shares common restriction sites with the 6.6-kilobase *Eco*RI fragment of the γ2b gene, the nucleotide sequence of which has already been determined[30]. Comparison of restriction maps of clone 2 and the γ2b clone indicates that clone 2 contains the γ2b gene at the 3′ end (Fig. 1).

However, restriction cleavage site mapping indicates that clone 9 does not contain any region corresponding to the γ2b structural gene (Fig. 1), but has one *Eco*RI fragment (5.2 kilobase) that hybridizes with the γ2b gene fragment (probe *d*) as shown in Fig. 2*A*, lane 2. Because the γ2b and γ2a genes are known to share extensive homology and cross-hybridize[30,35,36], we thought that clones 9 and 32 could contain the γ2a gene.

To test this we used as probe a γ2a cDNA clone pG2a–14 (probe *e* in Fig. 1) which does not hybridize with the γ2b gene in the stringent conditions used[17]. As shown in Fig. 2*B*, clone 9 was shown to contain an *Eco*RI fragment (5.2 kilobases) that hybridized to the γ2a gene probe (probe *e*). This *Eco*RI fragment is located at the 3′ end of the clone 9 insert and is cleaved by *Xho*I (data not shown). We have determined a partial nucleotide sequence of the 5.2-kilobase *Eco*RI fragment of clone 9. The nucleotide sequence from the *Bam*HI site towards the *Xho*I site of the γ2a coding region is

GGATCCCTGTCCAGTGGTGTG

CACACCTTCCCAGCTGTCCTGCAG,

which coincides with that reported for the CH1 domain of the γ2a cDNA clone[36]. The amino acid sequence

Gly-Ser-Leu-Ser-Ser-Gly-Val-His-

Thr-Phe-Pro-Ala-Val-Leu-Gln

predicted from the nucleotide sequence agrees well with that previously determined for the CH1 domain of the γ2a chain of MOPC173 (ref. 37). The sequence determines the direction of transcription of the γ2a gene and unequivocally demonstrates that clone 9 contains the γ2a gene at the 3′ side.

The restriction map analysis indicates that clones 9 and 32 are almost identical to each other except that clone 32 extends 1.7 kilobases further in the 5′ direction and has a small deletion (0.5 kilobases) in the middle (Fig. 1). It is clear that clone 32 has a γ2a gene fragment at the 3′ end identical to that of clone 9. In addition, clone 32 contains, at the 5′ end, a 3.6-kilobase *Eco*RI fragment which hybridizes with the γ2b gene probe (Fig. 2*A*, lane 3). The restriction map of the 5′-terminal portion (~2.6 kilobases) of clone 32 is indistinguishable from that of the 3′ end of clone 2 (Fig. 1). These results indicate that clone 32 shares overlapped regions at the 5′ end with clone 2 and at the 3′ end with clone 9. The linear alignment of clones 2, 32 and 9 clearly demonstrates that the γ2b gene is located 5′ to the γ2a gene and the two genes are 15 kilobases apart (Fig. 1).

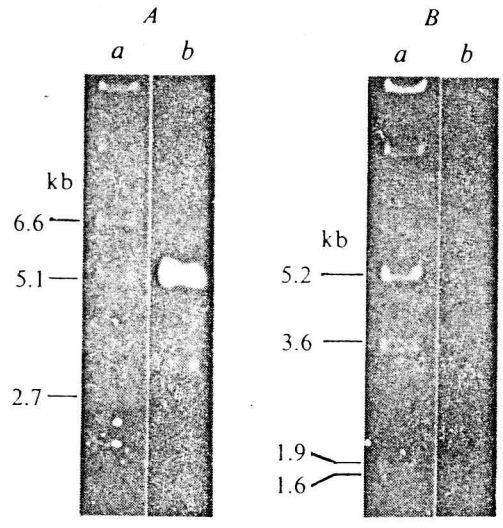


Fig. 3 Cross-hybridization between $S_{\gamma 1}$ and $S_{\gamma 2a}$ regions. Experiments were done as described in Fig. 2 legend. DNA was digested with *Eco*RI and Southern blot hybridized with ^{32}P-labelled 5′ *Eco*RI fragment (5.1 kilobases) of clone 3. Ethidium bromide stains(*a*) and autoradiograms(*b*) of Southern blots are shown. *A*, clone 3 DNA. *B*, clone 9 DNA. A faint band of 3.1 kilobases in lane *A*(*b*) is due to deletion of the 5.1-kilobase fragment of clone 3 during propagation.

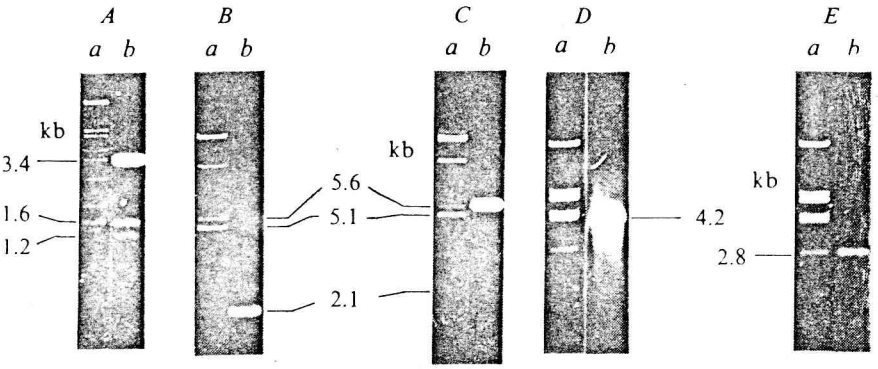


Fig. 4 Southern hybridization of isolated clones linking between the γ1 and γ2b genes and between the γ2a and ε genes. Experiments were done as described in Fig. 2 legend. Ethidium bromide stains(*a*) and autoradiograms(*b*) of Southern blots are shown. In *A*, clone 3 DNA was digested with *Eco*RI and *Bam*HI, and the 6.6-kilobase *Eco*RI fragment of the γ1 gene (Fig. 1*a*) was used as a probe. In *B*, clone 69 DNA was digested with *Eco*RI and 4.0-kilobase *Xba*I–*Hha*I fragment of the γ2b gene (Fig. 1*d*) was used as a probe. In *C*, clone 69 DNA was digested with *Eco*RI and 2.7-kilobase 3′ *Eco*RI fragment of clone 3 (Fig. 1*b*) was used as a probe. In *D*, clone 12 DNA was digested with *Eco*RI and *Bam*HI, and the probe was 2.2-kilobase *Bam*HI–*Hin*dIII fragment of the ε gene (Fig. 1*h*). *E*, clone 12 DNA was digested with *Eco*RI and *Bam*HI and the probe was the 5.2-kilobase *Eco*RI fragment of clone 9 (Fig. 1*f*).

The γ1 gene is located 5′ to the γ2b gene

The embryonic mouse DNA library was screened with a 6.6-kilobase γ1 gene fragment (IgH2) cloned previously[27] as a probe (probe *a* in Fig. 1) and two different clones were obtained: Ch · M · Igγ1 – 3 (clone 3) and Ch · M · Igγ1 – 17 (clone 17). Restriction maps of these are shown in Fig. 1.

Comparison of the restriction maps of clones 3 and 17, and the γ1 gene clone (IgH2) indicates that both clones 3 and 17 contain the 6.6-kilobase *Eco*RI fragment of the γ1 gene. The comparison also allows determination of the orientation of the two clones. The insert of clone 17 is ~2 kilobases shorter than that of clone 3 and the deletion is located in the 5′-terminal *Eco*RI fragment. These two clones may be derived from a common ancestral phage and the deletion may have occurred during propagation of phages, because the directions of the inserts within these phages and the ends of the inserts in these clones are identical to each other. Furthermore, clone 3 underwent deletion during propagation, producing a phage similar to clone 17 (see Fig. 3*A*). Southern blot experiments (Fig. 4*A*) provide additional evidence that clone 3 contains the γ1 gene. When clone 3 DNA was digested with *Eco*RI and *Bam*HI simultaneously, clone 3 DNA produced 3.4-, 1.6- and 1.2-kilobase fragments which hybridize to the γ1 gene (probe *a*), as expected from the restriction map.

Because γ1 gene clones (3 and 17) and γ2b gene clone (2) do not share a common region, we screened the embryonic mouse DNA library with a 3′ fragment of clone 3 (Fig. 1, probe *b*) and a 5′ fragment of clone 2 (Fig. 1, probe *c*) as probes. We isolated a new clone designated Ch · M · Igγ2b–69 (clone 69)—the restriction map of this is shown in Fig. 1.

The 3′ part of clone 69 is indistinguishable from that of clone 2 except that the former has an ~1.5-kilobase deletion at the 5′ side of the γ2b gene. Comparison of restriction maps of clones 69 and 2 indicates that clone 69 contains the γ2b gene at the 3′ end. Southern blot hybridization experiments show that a 2.1-kilobase *Eco*RI fragment of clone 69 hybridizes with the γ2b gene probe (Fig. 4*B*). Clone 69 also contains a 5.6-kilobase *Eco*RI fragment that hybridizes with the 3′-terminal fragment of clone 3 (Fig. 1, probe *b*) as shown in Fig. 4*C*. From these results it is clear that the 3′ region (2.1 kilobases) of clone 3 and the 5′ region (2.1 kilobases) of clone 69 are identical. The linear alignment of clones 3, 17, 69 and 2 clearly demonstrates that the γ1 gene is located 5′ to the γ2b gene and the two genes are 21 kilobases apart. These results are consistent with those of Maki *et al.*[20].

The ε gene is located 3′ to the γ2a gene

We recently cloned a γ2a gene fragment from *Eco*RI digests of RPC5 myeloma DNA designated Ch · M · Igγ2a–11 (clone 11). Southern blot hybridization and R-loop mapping show that the γ2a gene is located at the 5′ end of the 20-kilobase cloned fragment (data not shown). The restriction map of clone 11 is shown in Fig. 1. The 5′ region of the RPC5 γ2a gene fragment shares common restriction sites with the 3′ region of clones 32 and 9.

The 8.5-kilobase *Hin*dIII fragment (probe *g*) of clone 11 was recloned into pBR322 and used as a probe to screen the mouse embryonic DNA library. We obtained a germ-line γ2a gene fragment designated Ch · M · Igε–12 (clone 12)—the restriction map of this is also shown in Fig. 1. The 16.5-kilobase insert of clone 12 has restriction sites almost identical to those of the 3′ portion of clone 11 except that the latter has a deletion at the 3′ part.

Cloning of an expressed ε gene from an ε chain-producing hybridoma has been reported elsewhere[34]. Partial nucleotide sequence determination indicates that our ε gene clone Ch · M · Igε–1 encodes an amino acid sequence similar to that of human ε chain. The expressed ε gene (Ch · M · Igε–1) contains, as do other expressed H chain genes[20–26], at least three germ-line gene segments: a V gene, a J gene (including S_μ region that flanks the 5′ end of the μ gene) and the ε gene. The restriction map of the region encoding the ε gene of Ch · M · Igε–1, shown at the top of Fig. 1, seems to be identical to that of the 3′ region of clone 12.

To test whether or not clone 12 contains the ε gene, Southern blot hybridization experiments were done using a *Bam*HI–*Hin*dIII fragment (probe *h*) of the ε gene as a probe. Figure 4*D* shows that probe ε hybridizes a 4.2-kilobase *Bam*HI + *Eco*RI fragment at the 3′ part of clone 12. We also confirmed (Fig. 4*E*) that the 5′ segment of clone 12 has a region identical to the 3′ segment (probe *f*) of clone 9 which contains the germ-line γ2a gene. Taken together, these results indicate that the ε gene is located 3′ to the γ2a gene and the two genes are about 15 kilobases apart.

We have recently isolated three clones[38] which link the γ1 gene (clone 3) and the γ3 gene (Igγ3–30)[33]. Our data indicate that the γ3 gene is located 34 kilobases 5′ to the γ1 gene, and we

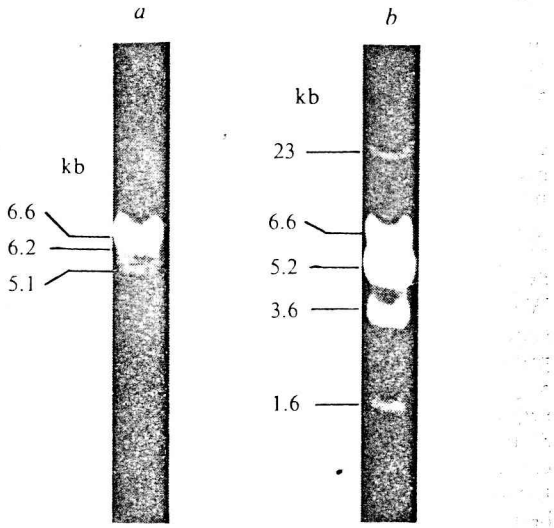


Fig. 5 Distance between the γ1, γ2b and γ2a genes. BALB/c newborn mouse DNA (3 μg) was digested with *Eco*RI to completion and run on 0.5% agarose gel electrophoresis. DNA was transferred to a nitrocellulose filter[19] and hybridized with ^{32}P-labelled probes as described elsewhere[17]. Total DNAs of clones 69 and 9, respectively, were used as probes in *a* and *b*.

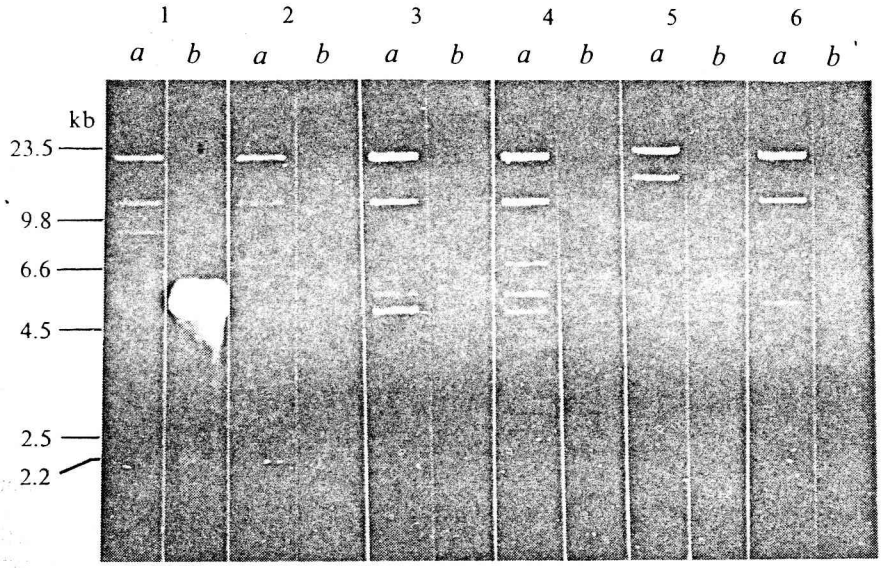


Fig. 6 Hybridization analysis with J-region probe. DNAs from isolated clones covering the region between the γ1 and γ2a genes were digested with *Eco*RI and run on 0.7% agarose gel electrophoresis. DNAs were transferred to a nitrocellulose filter[19] and hybridized with the 1.5-kilobase *Hin*dIII-*Eco*RI fragment that is located 6.5 kilobases 5′ to the μ gene. This fragment contains the J_{H4} region[24]. Ethidium bromide stains(*a*) and autoradiograms(*b*) of Southern blots are shown. DNA in each lane is as follows: (1) the expressed γ1 gene of MC101 myeloma (Igγ1-706)[25,47,48] that contains V, D, J_3, J_4, S_μ and $C_{\gamma 1}$ gene segments; (2) clone 3; (3) clone 69; (4) clone 2; (5) the 6.6-kilobase γ2b gene clone (IgH22)[28,30]; (6) clone 9.

have demonstrated[44] that the ε gene is present 12.5 kilobases upstream of the α gene. The linkage of J_H–μ–δ has been shown elsewhere[24,32]. Thus it is now clear that the organization of the mouse C_H gene is 5′-J_H–(6.5 kilobases)–μ–(4.5 kilobases)–δ–(unknown distance)–γ3–(34 kilobases)–γ1–(21 kilobases)–γ2b–(15 kilobases)–γ2a–(14.5 kilobases)–ε–(12.5 kilobases)–α–3′. Note that all the C_H genes ordered have the same orientation and thus are transcribed from the same strand of DNA.

Distance between C_H genes in germ-line DNA

Comparison of the restriction maps of the clones isolated indicate that some have deletions which probably occurred during propagation of phages. This was also reported for globin gene clones[39]. We therefore compared the sizes of *Eco*RI fragments obtained from cloned γ genes and from Southern blot hybridization of newborn mouse DNA. When the total insert of clone 69 was used as a probe, *Eco*RI digests of newborn mouse DNA produced 6.6(doublet)-, 6.2- and 5.1-kilobase bands which coincide with the *Eco*RI fragments deduced from the cloned DNAs (Fig. 5). When the total insert of clone 9 was used as a probe, *Eco*RI digests of newborn mouse DNA produced 23-, 6.6-, 5.2-, 3.6- and 1.6-kilobase bands which coincide with the *Eco*RI fragments deduced from the cloned DNAs.

Furthermore, consideration of the maximum size of the insert (21 kilobases) that can be taken by Charon 4A phage makes it unlikely that there is a deletion >2 kilobases in the clones obtained. Because the insert sizes of clones 69 and 32 are 19.3 and 19.2 kilobases, respectively, the original inserts cannot be more than 2 kilobases greater than those of the present clones. We therefore conclude that the distances between the γ1 and γ2b genes and between the γ2b and γ2a genes in the germ-line chromosome are 21.2 and 15.5 kilobases, respectively. Similar experiments using the ε gene as a probe[34] indicate that the distance between the γ2a and ε gene is 14.5 kilobases in newborn mouse DNA.

Absence of J region sequences in the entire region encompassing γ and ε genes

We determined whether or not each γ gene clone contains a sequence homologous to a J region. We used a J segment (J_4) as a probe and hybridized it with restricted DNA fragments of all the γ gene clones shown in Fig. 1. There were no J or J-like sequences between 8 kilobases 5′ to the γ1 gene and 3 kilobases 3′ to the γ2a gene (Fig. 6). The J_4 sequence was actually expressed in the γ2b chain produced by MOPC141 myeloma[24]. It can be argued that other completely different J sequences may be present in the γ gene-flanking regions. If this is the case, V region sequences formed by the use of J_μ and J_γ are different. This is unlikely because the same idiotype (a V region sequence) can be expressed as either μ or γ chain. It is possible that a γ gene has completely different J-flanking regions and an identical J_4 sequence of 51 base pairs which was not detected by our probe. This also seems unlikely because the precise joining between the same V gene and J_γ probably requires a nucleotide sequence similar to that of the J_μ-flanking region[23,24]. A similar experiment was done on the ε gene clone (clone 12) and a γ3 gene clone[33], in which we did not find a J region or J-like sequences (data not shown).

The S region before each C_H gene may comprise repetitive sequences

The class-switch recombination takes place at the region 5′ to each C_H gene. Such regions responsible for the class switch have been functionally defined as S regions[21]. We have recently demonstrated that the $S_{\gamma 1}$, $S_{\gamma 2b}$ and $S_{\gamma 3}$ regions comprise tandem repetitive conserved sequences of 49 base pairs[25,33]. The 5′-flanking region of the γ2b gene which contains the repetitive sequence is deleted from clone 69. Such tandem repetitive sequences of conserved 49-base pair units can be easily deleted when propagated in bacteria[40]. Deletions of short segments (0.5–2.0 kilobases) found in several clones are always located in the 5′-flanking region of each γ gene. The distance from the 5′ end of each γ gene to a deleted region is almost constant (~3 kilobases). During propagation of clone 9, we also found a deletion of ~1.5 kilobases at ~3 kilobases 5′ of the γ2a gene, about the same place where deletion occurred in clone 32. The deletions in these clones may indicate the positions of repetitive sequences.

Although we have not determined the nucleotide sequence of the $S_{\gamma 2a}$ region, the deletion in clone 32 may suggest a similar repetitive sequence in the 5′-flanking region of the γ2a gene. We found that the 5′-flanking region of the γ2a gene cross-hybridizes with the $S_{\gamma 1}$ region (Fig. 3). The S_μ and S_α regions also contain repetitive sequences[22,31,41] and these are often deleted during cloning. Thus it is likely that each C_H gene, except for the δ gene which is probably expressed by RNA splicing[32], has its own S region for class-switch recombination and that all the S regions comprise repetitive sequences.

We thank Dr P. Leder for an embryonic mouse DNA library, Dr K. Matsubara for critical reading of the manuscript, Mrs S. Nishida for technical assistance, and Miss T. Wada for preparation of the manuscript. Research was supported in part by grants from the Ministry of Education, Science and Culture of Japan and from Asahi Scientific Research Fund.

Received 30 September; accepted 24 November 1980.

1. Landucci-Tosi, S., Mage, R. G. & Dubiski, S. *J. Immun.* **104**, 641–647 (1970).
2. Kindt, T. J., Mandy, W. J. & Todd, C. W. *Biochemistry* **9**, 2028–2032 (1970).
3. Knight, K. L. & Hanly, W. C. *Contemp. Topics molec. Immun.* **4**, 55–88 (1975).
4. Nossal, G. J. V., Warner, N. L. & Lewis, H. *Cell Immun.* **2**, 41–53 (1971).
5. Fu, S. M., Winchester, R. J. & Kunkel, H. G. *J. Immun.* **114**, 250–252 (1975).

6. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1707–1711 (1975).
7. Goding, J. W. & Layton, J. E. *J. exp. Med.* **144,** 852–857 (1976).
8. Sledge, C., Fair, D. S., Black, B., Krueger, R. G. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **73,** 923–927 (1976).
9. Bleux, C., Ventura, M. & Liacopoulos, P. *Nature* **267,** 709–711 (1977).
10. Wang, A.-C., Wang, I. Y. & Fudenberg, H. H. *J. biol. Chem.* **252,** 7192–7199 (1977).
11. Kincade, P. W., Lawton, A. R., Bockman, D. E. & Cooper, M. D. *Proc. natn. Acad. Sci. U.S.A.* **67,** 1918–1925 (1970).
12. Honjo, T. & Kataoka, T. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2140–2144 (1978).
13. Rabbitts, T. H., Forster, A., Dunnick, W. & Bentley, D. L. *Nature* **283,** 351–356 (1980).
14. Colecliough, C., Cooper, D. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1422–1426 (1980).
15. Cory, S. & Adams, J. M. *Cell* **19,** 37–51 (1980.)
16. Cory, S., Jackson, J. & Adams, M. *Nature* **285,** 450–456 (1980).
17. Yaoita, Y. & Honjo, T. *Biomed. Res.* **1,** 164–175 (1980).
18. Yaoita, Y. & Honjo, T. *Nature* **286,** 850–853 (1980).
19. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
20. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 2138–2142 (1980).
21. Kataoka, T., Kawakami, T., Takahashi, N. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **77,** 919–923 (1980).
22. Takahashi, N., Kataoka, T. & Honjo, T. *Gene* **11,** 117–127 (1980).
23. Early, P. W., Huang, H., Davis, M. M., Calame, K. & Hood, L. *Cell* **19,** 981–992 (1980).
24. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286,** 676–683 (1980).
25. Honjo, T. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
26. Davis, M. H. *et al. Nature* **283,** 733–739 (1980).
27. Honjo, T. *et al. Cell* **18,** 559–568 (1979).
28. Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4240–4244 (1979).
29. Tucker, P. W., Marcu, K. B., Newell, N., Richard, J. & Blattner, F. R. *Science* **206,** 1301–1306 (1979).
30. Yamawaki-Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M. & Honjo, T. *Nature* **283,** 786–789 (1980).
31. Kawakami, T., Takahashi, N. & Honjo, T. *Nucleic Acids Res.* **8,** 3933–3945 (1980).
32. Liu, C.-P., Tucker, P. W., Mushinski, J. F. & Blattner, F. R. *Science* **209,** 1348–1353 (1980).
33. Kataoka, T., Miyata, T. & Honjo, T. *Cell* (in the press).
34. Nishida, Y. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
35. Yamawaki-Kataoka, Y. *et al. Biochemistry* **18,** 490–494 (1979).
36. Sikorav, J.-L., Auffray, C. & Rougeon, F. *Nucleic Acids Res.* **8,** 3143–3155 (1980).
37. Fougereau, M., Bourgois, A., de Preval, C., Rocca-Serra, J. & Schiff, C. *Annls Immun. Inst. Pasteur, Paris* **127,** 607–631 (1976).
38. Takahashi, N., Shimizu, A. & Honjo, T. (in preparation).
39. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Cell* **19,** 959–972 (1980).
40. Tsujimoto, Y. & Suzuki, Y. *Cell* **16,** 425–436 (1979).
41. Davis, M. M., Kim, S. K. & Hood, L. E. *Science* **209,** 1360–1365 (1980).
42. Maniatis, T. *et al. Cell* **15,** 687–701 (1978).
43. Blattner, F. R. *et al. Science* **196,** 161–169 (1977).
44. Benton, W. D. & Davis, R. W. *Science* **196,** 180–182 (1977).
45. Maniatis, T., Jeffrey, A. & Kleid, D. K. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1184–1188 (1975).
46. Tiemeier, D. C., Tilghman, S. M. & Leder, P. *Gene* **2,** 173–191 (1977).
47. Obata, M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
48. Kataoka, T. *et al.* (in preparation).

Reprinted from Nature, Vol. 283, No. 5749, pp. 786-789, February 21 1980

Complete nucleotide sequence of immunoglobulin γ2b chain gene cloned from newborn mouse DNA

Yuriko Yamawaki-Kataoka, Tohru Kataoka, Naoki Takahashi, Masahiro Obata & Tasuku Honjo*

Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo 113, Japan

Immunoglobulin heavy chain protein comprises five classes μ, γ, α, δ and ε chains which differ in amino acid sequences of the constant (C) region. The γ heavy chain class, which contains three domains and the hinge region, is divided into several subclasses such as γ1, γ2a, γ2b and γ3 in mouse. Amino acid sequence studies of mouse and human heavy chain proteins have shown that the γ subclass heavy chains are more closely related to each other than to the heavy chains of other classes. A similar conclusion was drawn from a nucleic acid hybridisation study using cDNAs complementary to the purified heavy chain mRNAs of the γ subclasses[1]. These results suggest that the immunoglobulin heavy chain genes of the γ subclasses have evolved by relatively recent gene duplication of the ancestral γchain gene. The complete nucleotide sequence of the γ1 chain gene has demonstrated unequivocally that the domains and the hinge region are encoded in separate segments of DNA[2]. Partial nucleotide sequences of the γ2b chain gene indicated that the γ2b gene is interrupted by intervening sequences (IS) at locations homologous to those in the γ1 chain gene[3]. Comparison of the structures of these closely related genes would give new insights into evolution of eukaryotic genes and serve for understanding biological significance of various segments of the genes. We have now determined the complete nucleotide sequence of the immunoglobulin γ2b chain gene cloned from newborn mouse DNA. The sequence not only predicts the complete amino acid sequence of the γ2b chain but also demonstrates that the γ2b chain gene is interrupted by three intervening sequences at the junction of the domains and the hinge region.

The 6.6-kilobase *Eco*RI fragment coding for the C region of the immunoglobulin γ2b chain was previously cloned from newborn mouse DNA using λgtWES·λB as an EK2 vector[3]. The recombinant phage was referred to as λgtWES·IgH22 (abbreviated IgH22). The 6.6-kilobase fragment was isolated by agarose gel electrophoresis and ligated with the *Eco*RI-digested pBR322[4]. The hybrid plasmid was referred to as pBR322·IgH22 (abbreviated pIgH22). The detailed restriction enzyme cleavage map of the *Eco*RI–*Hha*I fragment was constructed and the nucleotide sequence was determined by the chemical modification method of Maxam and Gilbert[5]. The restriction enzyme cleavage sites used for sequencing and the direction and range of sequencing are shown in Fig. 1.

The complete nucleotide sequence of the γ2b chain C region gene is shown in Fig. 2. The total number of the nucleotides determined is 1,834, which include the coding sequence, the IS, the 3′ untranslated sequence and the 5′ and 3′ flanking sequences.

* To whom correspondence should be addressed. Present address: Department of Genetics, Osaka University School of Medicine, Kita-ku, Osaka 530, Japan.

The nucleotide sequence predicts the complete amino acid sequence of the C region of the γ2b chain (336 residues) which consists of the C_{H1} (residue 118–214), C_{H2} (residue 237–346) and C_{H3} domains (residue 347–453) and the hinge region (residue 215–236), respectively. The predicted sequence was compared with the partial amino acid sequences so far published, including 18 residues in the hinge region and the C_{H2} domain[6], 66 residues in the C_{H2} domain and 35 residues in the C_{H3} domain[7] (Fig. 2). The amino acid sequences derived from the nucleotide sequences matched the known amino acid sequences except two residues, Ser in place of Leu at residue 240 and Ser in place of Tyr at residue 268. To determine the junctions of the IS and coding sequences (see below) the nucleotide sequences were also correlated with the amino acid sequences of the γ2a chain[7,8] which share extensive homology with the known amino acid sequences of the γ2b chain[7].

The complete amino acid sequence of the C region of the γ2b chain is compared with the amino acid sequences of the γ1[2] and γ2a[7,8] chains as shown in Fig. 3. It is evident that the γ2b and γ2a chains show more homology to each other than to the γ1 chain. The results agree with our previous conclusion from cDNA–mRNA hybridisation studies[1]. The homology between the three γ chain subclasses (γ1, γ2a and γ2b) for each domain is calculated. The C_{H1} domains are most conserved (79–82%) throughout the three subclasses while the C_{H3} domains are least conserved (53–59%). The C_{H2} domains are highly conserved between the γ2a and γ2b chains whereas they are moderately conserved (63–68%) between the γ1 and γ2b (or γ2a) chains. The G + C content of the coding sequences is 52.3%, exceeding the value (41.8%) of the total mouse DNA[9].

Note that the nucleotide sequences predict an unexpected Lys residue at the carboxy terminus of the γ2b chain. Although most of the γ chains so far sequenced end with the Pro-Glys residues[10], the nucleotide sequences of the cDNA as well as chromosomal gene of the γ1 chain also predict an additional Lys residue at the end of the heavy chain[2,11]. These results suggest that the newly-synthesised γ2b chain proteins may have a Lys residue at the carboxy terminus, which will be removed eventually either at secretion or during circulation in the blood. It would be interesting to see whether the γ1 and γ2b chain proteins bound to the cellular surface have the Lys residue at their carboxy terminus since the amino acid sequencing was done using secreted immunoglobulins. The membrane-bound and secreted μ chains are reported to be different in the primary structures of their carboxy termini[12,13].

The comparison of the nucleotide sequences of the γ2b chain gene with the amino acid sequences of the γ2b and γ2a chains indicates that the domains and the hinge regions of the γ2b chain structural sequence are separated by three ISs as shown in Figs 1 and 2. The ISs split the γ2b chain gene at positions homologous to those in the γ1 chain gene[2,14]. The first, second and third ISs of the γ1 and γ2b chain genes as a pair are 356/316, 98/107, 120/112 base pairs long, respectively. The evolutionary significance of the ISs in immunoglobulin genes was discussed more extensively in previous reports[2,3].

The boundary sequences are compared between the γ1 and γ2b chain genes (Fig. 4). The first 10 nucleotides at each end of the ISs are highly conserved, although not identical, between the two genes whereas there is considerable divergence of the overall nucleotide sequences of the homologous IS. The possibility is discussed that not only nucleotide sequences at the joints of the IS and structural sequences but also other signals such as

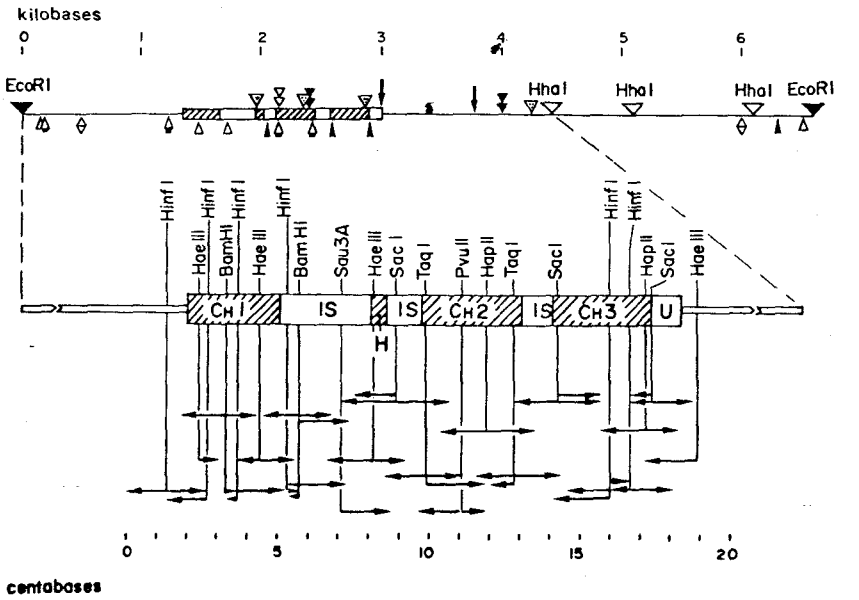


Fig. 1 Sequencing strategy of the mouse γ2b chain constant region gene. Plasmid DNA pIgH22 was purified according to the method of Clewell and Helinski[16]. Restriction cleavage sites were determined by combined cleavages of two or more enzymes. *Bam*HI, *Sac*I, *Xho*I and *Kpn*I sites were determined previously[3]. The 4.4-kilobase *Eco*RI–*Hha*I fragment was isolated by 4% polyacrylamide gel electrophoresis after digestion of pIgH22 with *Eco*RI and *Hha*I, cleaved with other restriction enzymes and used as the substrates of phosphorylation. Alternatively, IgH22 was digested with either *Bam*HI or *Sac*I and resulting fragments were isolated. Each fragment was digested with the second restriction enzymes and used for phosphorylation. Sources of restriction enzymes were described previously[2]. The length of sequenced fragment is given in centabases in the 5' to 3' direction relative to the mRNA sequence. The horizontal arrows shown at the bottom indicate the direction and the extent of the sequence determined. The restriction map of the cloned 6.6-kilobase *Eco*RI fragment is presented on the top line. The restriction sites which were used for sequencing in the 4.4-kilobase *Eco*RI–*Hha*I fragment are shown on the rectangle below. The γ2b chain gene is represented by the wide rectangle. C_{H1}, C_{H2}, C_{H3}, domains of constant region; H, hinge region; IS, intervening sequence; U, untranslated sequence.

```
                    20                  40                  60                  80
0   GGAAATTTCTGACACTGCTTGCCTCAGATCAATTGTAGAAGACACGGTTCTAAGAGCAAAGCTAAGAACAGAATCTCTCCAAATATCCGAGGCCACTGAT

                                                A  K  T  T  P  P  S  V  Y  P  L  A  P  G  C  G  D  T  T  G  S  S  V
100 AAGAAAAAGCTCACACATTCTCCTCTCTTGCAGCCAAAACAACACCCCCATCAGTCTATCCACTGGCCCCTGGGTGTGGAGATACAACTGGTTCCTCCGT

                  T  L  G  C  L  V  K  G  Y  F  P  E  S  V  T  V  T  W  N  S  G  S  L  S  S  S  V  H  T  F  P  A  L
200 GACTCTGGGATGCCTGGTCAAGGGCTACTTCCCTGAGTCAGTGACTGTGACTTGGAACTCTGGATCCCTGTCCAGCAGTGTGCACACCTTCCCAGCTCTC

                  L  Q  S  G  L  Y  T  M  S  S  S  V  T  V  P  S  S  T  W  P  S  Q  T  V  T  C  S  V  A  H  P  A  S
300 CTGCAGTCTGGACTCTACACTATGAGCAGCTCAGTGACTGTCCCCTCCAGCACTTGGCCAAGTCAGACCGTCACCTGCAGCGTTGCTCACCCAGCCAGCA

                  S  T  T  V  D  K  K  L
400 GCACCACGGTGGACAAAAAACTTGGTGAGAGGACATTCAGGGGAGGAGGGATTCACCAGAGTTGAGGCAAAGGTATTAGCCTATCTAAACCAGCCAGGCT

500 GGGATCCATCACCAAGGAGGTGACCTTAGCCCAGGGAAGAGGGAGATACTGTCTCTGCCTCCCTCCTGGGAACATCTAGCTATGACCACCTACACTCAAG

600 GACATGTTCCTCTGGGATAGGTGTGCTTGTCATTTCCAGGATCATCCTGGAACTAAGCCCATACCAGGACAAACTTTCCTCTCTCTGGTTTGGTGCTTCT

                                                                    I  N  P  C  P  P  C  K  E  C  H  K
                                                E  P  S  G  P  I  S  T  I  N  P  C  P  P  C  K  E  C  H  K
700 CTCCTTCAAAAACCAGTAACATCCAGCCTTCTCTCTGCAGAGCCCAGCGGGCCCATTTCAACAATCAACCCCTGTCCTCCATGCAAGGAGTGTCACAAAT

                  C  P
                  C  P
800 GCCCAGGTAAGTCACTACCAGAGCTCCACTCCCAGGAGAATGGTAAGTGCTGTAAAAATCCCTGTAATGGAGGATAAGCCATGTACAAATCCATTTCCAT

                                A  P  N  L  E  G  G  P  L  V  F  I  F  P  P  N  I  K  D  V  L  M  I  S  L  T  P  K  V
                                A  P  N  L  E  G  G  P  S  V  F  I  F  P  P  N  I  K  D  V  L  M  I  S  L  T  P  K  V
900 CTCTCCTCATCAGCTCCTAACCTCGAGGGTGGACCATCCGTCTTCATCTTCCCTCCAAATATCAAGGATGTACTCATGATCTCCCTGACACCCAAGGTCA

                T  C  V  V  V  D  V  S  H  E  D  P  E  V  K  F  N  W  Y  V  D  G  V  E  V  H  N  A  K  T  K  P  R
1000 CGTGTGTGGTGGTGGATGTGAGCCACGAAGACCCAGACGTCCAGATTCAGCTGGTTTGTGAACAACGTGGAAGTACACACAGCTCAGACACAAACCCATAG

                                                                S  G  K  E  F  K  C  K  V  N  N  K
                E  D  Y  N  S  T  I  R  V  V  S  T  L  P  I  Q  H  Q  D  W  M  S  G  K  E  F  K  C  K  V  N  N  K
1100 AGAGGATTACAACAGTACTATCCGGGTGGTCAGCACCCTCCCCATCCAGCACCAGGACTGGATGAGTGGCAAGGAGTTCAAATGCAAGGTCAACAACAAA

                D  L  P  S  P  I  E  K  T  I  S  K  I  K
                D  L  P  S  P  I  E  R  T  I  S  K  I  K
1200 GACCTCCCATCACCCATCGAGAGAACCATCTCAAAAATTAAAGGTGGGACCTGCAGGACAACTGCATGGGGGCTGGGATGGGCATAAGAATAAATGTCTA

                                                                V
                                                G  L  V  R  A  P  Q  V  Y  I  L  P  P  P  A
1300 TGTGGACAGCCTTCCACTTCAGCCATGACCTCTATGTATGTTTCTAACCCCACAGGGCTAGTCAGAGCTCCACAAGTATACATCTTGCCGCCACCAGCAG

                E  Q  L  S  R  K  D  V  S  L  T  C  L  V  V  G  F  N  P  G  D  I  S  V  E  W  T  S  N  G  H  T  E  E
1400 AGCAGTTGTCCAGGAAAGATGTCAGTCTCACTTGCCTGGTCGTGGGCTTCAACCCTGGAGACATCAGTGTGGAGTGGACCAGCAATGGGCATACAGAGGA

                                                                K  T  S  K  W  E  K  T  D  S
                N  Y  K  D  T  A  P  V  L  D  S  D  G  S  Y  F  I  Y  S  K  L  N  M  K  T  S  K  W  E  K  T  D  S
1500 GAACTACAAGGACACCGCACCAGTCCTAGACTCTGACGGTTCTTACTTCATATATAGCAAGCTCAATATGAAAACAAGCAAGTGGGAGAAAACAGATTCC

                F  S  C  N  V  R  H  E  G  L  K  N  Y  Y  L  K  K  T  I  S  R  S  P  G
                F  S  C  N  V  R  H  E  G  L  K  N  Y  Y  L  K  K  T  I  S  R  S  P  G  X Term
1600 TTCTCATGCAACGTGAGACACGAGGGTCTGAAAAATTACTACCTGAAGAAGACCATCTCCCGGTCTCCGGGTAAATGAGCTCAGCACCCACAAAGCTCTC

                                                                *poly(A)
1700 AGGTCCTAAGAGACACTGGCACCCATATCCATGCATCCCTTGTATAAATAAAGCATCCAGCAAAGCCTGGTACCATGTAAAACTGTCCTGGTTCTTTCCA

1800 AGGTATAGAGCATAGCTCACGGGCTGATGGGTCT
```

Fig. 2 Nucleotide sequence of mouse γ2b chain gene. The nucleotide sequence of the anti-coding strand is displayed 5' to 3'. DNA sequencing was performed according to the method of Maxam and Gilbert[5]. DNA fragments to be sequenced were digested with the enzymes indicated in Fig. 1, treated with bacterial alkaline phosphatase (Boehringer) and phosphorylated using T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (NEN). ^{32}P-DNA fragments were isolated by polyacrylamide gel electrophoresis and cleaved with second restriction enzymes. Fragments labelled only at one end were isolated by polyacrylamide gel electrophoresis and purified as described previously[2]. For strand separation the labelled DNA samples were dissolved in 100 μl of 0.2 M KOH, 10% (w/v) glycerol, 0.01% (w/v) xylene cyanol and 0.01% (w/v) bromphenol blue, heated at 37 °C for 1 min and separated with 15% or 20% polyacrylamide gel. Four base-specific reactions were utilised (G, A>C, C+T, C). Samples were loaded in 90% formamide and separated with thin gels (0.5×300×400 mm, 8 and 20% polyacrylamide gel). The amino acids predicted by the nucleotide sequence are shown above the coding sequences in italics. The partial amino acid sequences of MOPC 141[6] and MPC 11[7] are presented on the top line. Term, termination codon; poly (A), putative poly (A) addition site assigned by comparison with the mouse γ1 chain gene sequence. Amino acids are expressed by one letter code of A(Ala), C(Cys), D(Asp), E(Glu), F(Phe), G(Gly), H(His), I(Ile), K(Lys), L(Leu), M(Met), N(Asn), P(Pro), Q(Gln), R(Arg), S(Ser), T(Thr), V(Val), W(Trp) and Y(Tyr).

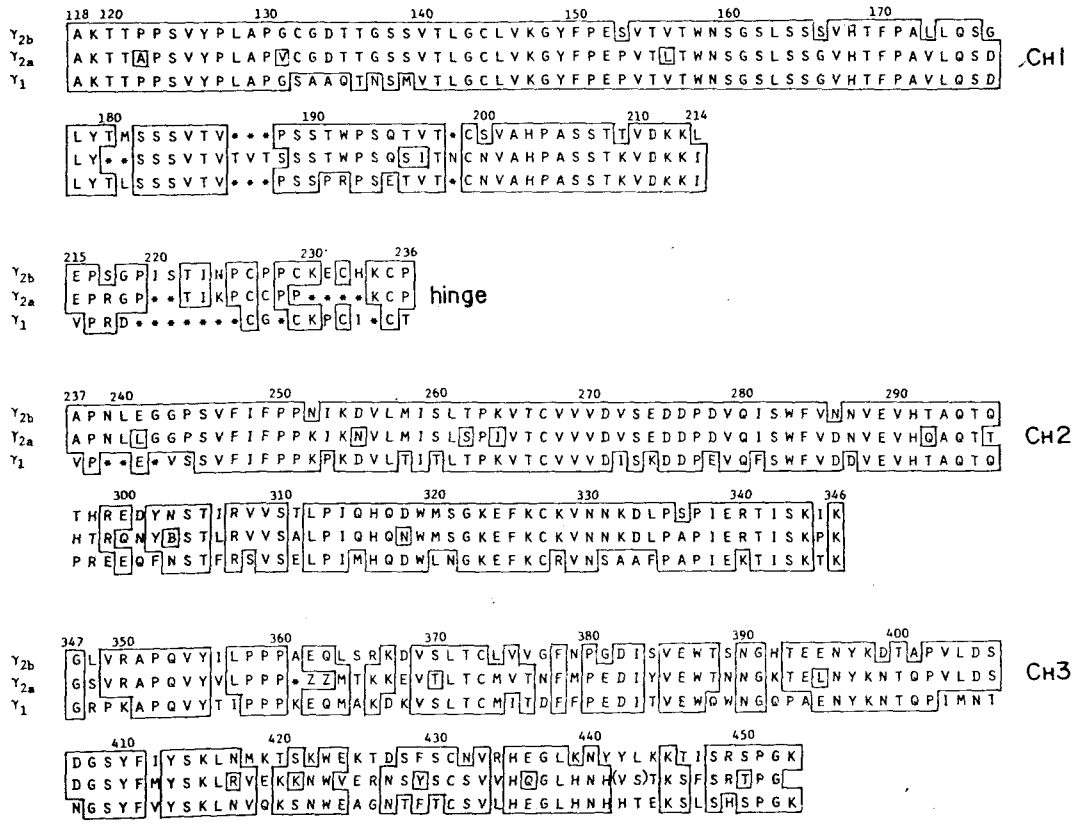


Fig. 3 Comparison of the amino acid sequences of $\gamma 1$, $\gamma 2a$ and $\gamma 2b$ chains. The amino acid sequences of the $\gamma 1$ and $\gamma 2b$ chains were predicted by the nucleotide sequences of the respective genes cloned from newborn mouse DNA[2]. The amino acid sequence of the $\gamma 2a$ chain was taken from the data of MOPC 173[7,8]. The residue was numbered with the first residue of C_{H1} domain as 118. Homologous positions are boxed. The one letter code used is as described in Fig. 2.

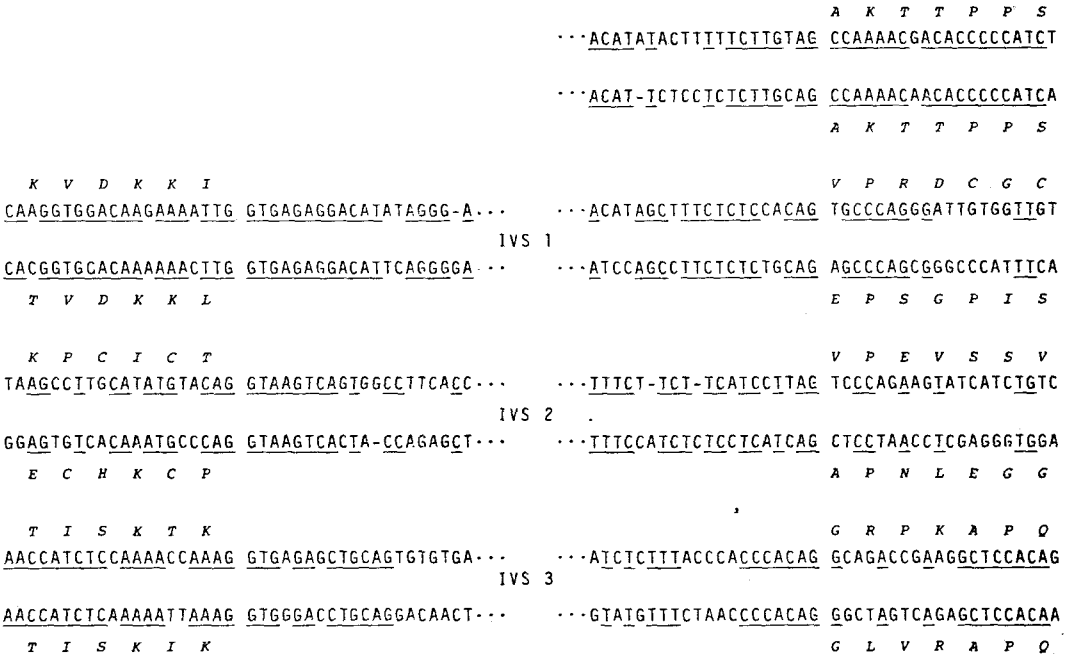


Fig. 4 Comparison of nucleotide sequences of the $\gamma 1$ and $\gamma 2b$ chain genes around the junction of IS and coding sequences. Twenty nucleotides each of the homologous IS and coding sequences of the $\gamma 1$ (upper sequences) and $\gamma 2b$ (lower sequences) chain genes are aligned to maximise homology. The top row is the 5′ end of the C_{H1} domain. The second, third and fourth rows are the 5′ and 3′ ends of the first, second and third IS, respectively. The nucleotide sequence of the strand corresponding to the mRNA is displayed 5′ to 3′ direction. The homologous nucleotides are underlined. Amino acids predicted from the nucleotide sequences are shown in italic letter by one letter code as described in Fig. 2.

secondary and tertiary structures of mRNA precursors may be important for the recognition of splicing[15].

We thank Dr Y. Mano for his encouragement. This investigation was supported in part by grants from the Ministry of Education, Science and Culture of Japan.

Received 1 October; accepted 30 November 1979.

1. Yamawaki-Kataoka, Y. *et al. Biochemistry* **18**, 490–494 (1979).
2. Honjo, T. *et al. Cell* **18**, 559–568 (1979).
3. Kataoka, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 4240–4244 (1979).
4. Bolivar, F. *et al. Gene* **2**, 95–113 (1977).
5. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
6. de Preval, C., Pink, J. R. L. & Milstein, C. *Nature* **228**, 930–932 (1970).
7. Francus, T. & Birshtein, B. K. *Biochemistry* **17**, 4324–4331 (1978).
8. Fougereau, M. *et al. Annls Immun.* **127C**, 607–631 (1976).
9. Maio, J. J. in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* **II**, 379–404 (CRC Press, Ohio, 1976).
10. Fudenberg, H. H., Pink, J. R. L., Stites, D. P. & Wang, A. C. in *Basic Immunogenetics*, 15 (Oxford University Press, New York, 1978).
11. Obata, M. *et al. Gene* (in the press).
12. Kehry, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 2932–2936 (1979).
13. Williams, P. B., Kubo, R. T. & Grey, H. M. *J. Immun.* **121**, 2435–2439 (1978).
14. Sakano, H. *et al. Nature* **277**, 627–633 (1979).
15. Obata, M. *et al.* (in preparation).
16. Clewell, D. & Helinski, D. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1159–1166 (1969).

100. *Comparison of Immunoglobulin γ1 and γ2b Chain Genes Cloned from Newborn Mouse*

Evolutional Divergence of Intervening Sequences and Structural Basis for Recognition of Splicing Sites

By Masahiro OBATA, Yuriko YAMAWAKI-KATAOKA, Naoki TAKAHASHI, Tohru KATAOKA, Toshiaki KAWAKAMI, and Tasuku HONJO*)

Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo 113

(Communicated by Osamu HAYAISHI, M. J. A., Dec. 12, 1979)

Accumulating evidence demonstrates that the coding sequences of most eukaryotic genes are interrupted by more than one intervening sequences (IVS). The IVS as well as coding sequences are shown to be transcribed as continuous primary nuclear RNA, followed by splicing of the coding sequence.

An important question is what is the structural basis for highly specific recognition of the splicing machinery. Comparisons of boundary sequences between different IVS of the ovalbumin and the mouse β-globin genes have revealed common dinucleotides GT at the 5′ end and AG at the 3′ end of the IVS.[1),2)] Although boundary sequences of the homologous IVS of mouse and rabbit β-globin genes share longer common nucleotides in general, the shortest of the common sequences is AG at the 3′ end of the small intervening sequence.[3)] Obviously such short sequences alone are not capable of serving strong specificity for splicing. It may be difficult to reveal common sequences responsible for splicing signals by comparing diverse genes or the same genes in different animals if different sets of the splicing enzymes catalyze processing of the primary transcripts of diverse genes. It is thus important to compare the structure of closely related genes in order to understand structural basis for splicing.

We have recently cloned two immunoglobulin heavy chain genes of the γ subclasses γ1 and γ2b from the newborn mouse[4),5)] and compared the nucleotide sequences of the homologous IVS of these closely related genes.

Comparison of the γ1 and γ2b chain genes by restriction mapping. The mouse immunoglobulin γ1 and γ2b chain DNA fragments purified

*) Present address: Department of Genetics, Osaka University School of Medicine, Kita-ku, Osaka 530, Japan.

from EcoRI digests of chromosomal DNA cloned in λgtWES were transposed into the EcoRI site of pBR322 by conventional technique. Detailed restriction maps around the structural sequences of IgH2 (γ1 gene) and IgH22 (γ2b gene) are compared as shown in Fig. 1.

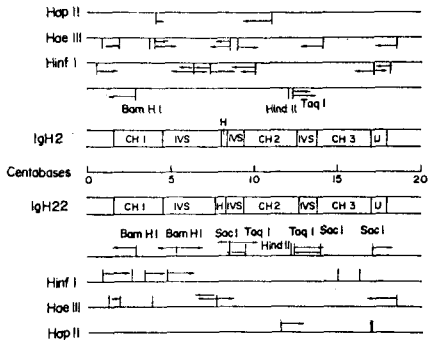


Fig. 1. Comparison of the restriction maps of the immunoglobulin γ1 and γ2b chain genes cloned from newborn mouse. The horizontal arrows indicate the direction and range of the sequence read for IVS regions. CH_1, CH_2, CH_3, domains; H, the hinge region; U, the 3' untranslated region.

Though most of the restriction sites shown are not conserved between two genes, a Bam HI site in the C_{H1} region, a Hind II site and a Taq I site in the C_{H2} region are common between the two genes.
between the two genes.

Locations and lengths of IVS in the γ1 and γ2b chain genes. The IVS split the coding sequences of the γ1 and γ2b chain genes virtually at the identical positions *i.e.* between the domains and the hinge region as shown in Fig. 1. Similar results were reported for the γ1 chain gene cloned from a mouse myeloma.[6] The α chain gene seems to be separated into similar segments although the exact locations of IVS are not known.[7] The result implicates that the IVS was introduced into the heavy chain gene before the divergence of heavy chain classes. The results also support the hypothesis that eukaryotic genes have been discontinuous from the beginning.[4)-6)]

Three IVS were numbered with the direction of transcription. The first, second and third IVS of the γ1 and γ2b chain genes as a pair are 356/317, 98/107 and 120/111 nucleotides long, respectively. The lengths of the homologous IVS are essentially similar between the two genes. Possibly a certain range of the length of IVS may be important for splicing though the precise length of IVS is not required. The larger IVS of two mouse β globin genes are also of similar length.[8]

Comparison of nucleotide sequences between the γ1 and γ2b chain

genes. The similarity index[3] can be used as a rough estimate of homology between two sequences. The index expresses complete and random match as 100 and 25, respectively. As shown in Table I, the

Table I. Comparison of nucleotide sequences of the γ1 and γ2b chain genes cloned from newborn mouse

Gene segment	Nucleotide compared (γ1/γ2b)	Gaps (bases)	Corrected length	Similarity index (%)
C_{H1} domain	291/291	21	301	86
Hinge region	39/66	36	70	41
C_{H2} domain	321/330	33	342	74
C_{H3} domain	321/321	53	350	70
3' untranslated region	175/156	29	180	66
IVS 1	356/317	87	381	64
IVS 2	98/107	39	122	60
IVS 3	120/111	39	136	59

Sequences are aligned to maximize homology by a computer using a modified program of Needleman and Wunsch.[24] Similarity index[3] is computed as 100−(% gap+% transition+% transversion). The percentage is calculated relative to corrected length which is obtained by adding the number of base pairs and the number of gaps of the aligned sequences.

similarity indexes of the homologous domains of the γ1 and γ2b chain genes are 70–86, showing extensive conservation of the structural sequences. The similarity index of the hinge regions shows an extensive divergence though it is certainly under-estimated because two sequences are of quite different length. The similarity indexes of the homologous IVS and the 3′ untranslated region are smaller than the value of the coding sequences but greater than the values (36–38) obtained by comparing two unrelated sequences, the β globin gene and the domains of the γ1 chain genes.[5] These results suggest that the homologous IVS of the γ1 and γ2b chain genes share common ancestors and have diverged considerably faster than the structural sequences during evolution.

Nucleotide sequences at the joint of the IVS and coding sequences. The nucleotide sequences around the boundaries of the coding sequences and IVS of IgH2 and IgH22 are shown in Fig. 2. All the IVS begin with GT and end at AG as previously indicated by Breathnach *et al.*[2] When the first 10 nucleotides each of the IVS and coding sequences at the homologous junctions of the γ1 and γ2b chain genes are compared, the IVS are at least 70% homologous to each other, which is not smaller than the values (30–90%) of the coding sequences. The results suggest that the nucleotide sequences at the extreme end

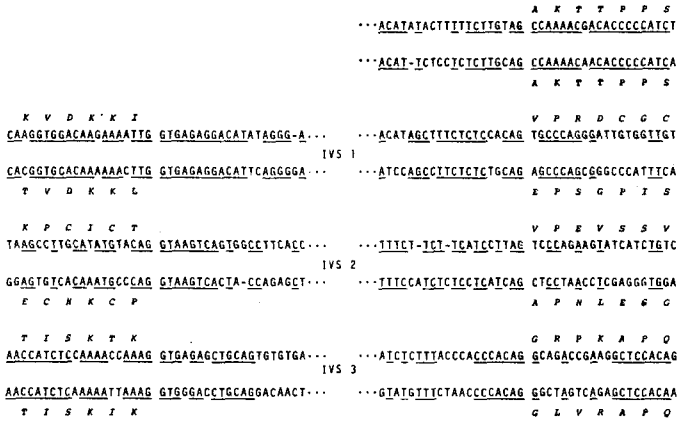


Fig. 2. Comparison of nucleotide sequences of the γ1 and γ2b chain genes around the junction of IVS and coding sequences. The nucleotide sequence determination was performed according to Maxam and Gilbert[23] as described elsewhere.[5] Twenty nucleotides each of the homologous IVS and coding sequences of IgH2 (upper sequences) and IgH22 (lower sequences) are aligned to maximize homology. The top row is the 5' end of the C_{H1} domain. The second, third and fourth rows are the 5' and 3' ends of the first, second and third IVS, respectively. The nucleotide sequence of the strand corresponding to the mRNA is displayed 5' to 3' direction. The homologous nucleotides are underlined. Amino acids predicted from the nucleotide sequences are shown in italic letter by one letter code.

of each junction are important for the recognition of definite splicing sites by the splicing machinery which excises and ligates precisely at the definite nucleotide positions, and thus considerably conserved between the two genes.

Even if there are multiple sets of the splicing enzymes, it is reasonable to assume that the primary transcripts of the γ1 and γ2b chain genes, which are expressed with a common set of the variable region genes in the differentiated bone marrow-derived lymphocyte, share a common set of the splicing enzymes. Otherwise we should expect that the total number of the mouse splicing enzyme set is as many as the total number of mouse genes. Nevertheless, we are unable to find a universal sequence at all the joints except for dinucleotides GT at 5' end and AG at 3' end of the IVS. Nor a long stretch of the nucleotide sequences common to both γ1 and γ2b chain genes are present around each splicing point. These results suggest that the splicing machinery recognizes not only nucleotide sequences at the joints of the IVS and structural sequences but also some other signals to specify the splicing points. The secondary or tertiary structure of RNA was discussed as a possible candidate for such recognition.[1),6),9)] Combination of short nucleotide sequence and secondary structure may serve a strong specificity as shown by processing of rRNA and

T7 phage mRNA by RNase III.[10)-12)]

Deletion mutants. There are mutant cell lines of mouse myeloma which synthesize altered heavy chains. P3(MOPC21, γ1) has yielded a spontaneous variant, IF2 that deletes the complete C_{H1} domain.[13)-14)] Mutagenesis of the MPC 11 (γ2b) mouse myeloma cell line has yielded a number of variants which synthesize heavy chains either shorter than the parent or of subclass different from the parent.[15)-16)] The former is of three quarters of length of the parent heavy chain. The latter is shown to have the V region identical to MPC 11 and the C region of the γ2a subclass.[17)] We suspect that the first group of these altered heavy chains contain the C region lacking a whole domain and the second group has the MPC 11 V region directly linked to the γ2a C region although the detailed chemical structure is not yet available.

Inasmuch as each domain is separated from other structural sequences by IVS, IF2 and variants of MPC 11 may have lost a whole domain gene sequence by recombination somewhere in IVS. Alternatively, point mutations at the essential region for the splicing signal have modified splicing frame so that the domains are trimed off from mRNA precursors. In any event the heavy chain genes in the variants mentioned above have mutated in such a way that the spliced pairs of the domains or the hinge region are changed. In view of the fact that there are no definite boundary nucleotide sequences conserved between two immunoglobulin genes, we prefer the former possibility. If so, the phenomena implicate that the spliced pair of the coding sequences is not determined precisely by the nucleotide sequence alone but different combinations are feasible as long as the topological order and/or other signals such as the secondary structure allows.

J segments. Recently Max *et al.*[18)] and Sakano *et al.*[19)] have shown that immunoglobulin κ light chain gene has five J segments with identical nucleotide sequences at the joints of the IVS and structural sequences. At least four of them seem to be capable of joining with the C region gene sequence by splicing. Seidman *et al.*[20)] demonstrated that in MOPC 41 myeloma the V region gene is linked to the J segment farthest from the Cκ region, leaving four more J segments intact between the V and the C region genes. When this gene is transcribed into the mRNA precursor several combinations of spliced pairs are possible between the J and Cκ region sequences. Yet, MOPC 41 κ chain mRNA and the protein synthesized are homogeneous,[21)-22)] suggesting that only one combination is preferred. These results also support our conclusion that the recognition of splicing sites depends on not only the nucleotide sequences at the boundary of the structural sequences and IVS but also other structural signals such as secondary and tertiary structures of RNA.

Acknowledgements. We are grateful to Drs. Y. Suzuki and M. Obinata for critical readings of the manuscript, to Dr. T. Miyata for providing his computer program and to Dr. Y. Mano for continuous encouragement. This investigation was supported by the grants from Ministry of Education, Science and Culture. Cloning experiments were done in a P2 facility using EK2 biological containment to meet Japanese Guideline for Recombinant DNA Research described in Japanese Scientific Monthly, vol. 31, p. 626 (1978).

References

1) Konkel, D. A., Tilghman, S. M., and Leder, P.: Cell, **15**, 1125 (1978).
2) Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., and Chambon, P.: Proc. Natl. Acad. Sci. U.S.A., **74**, 4853 (1978).
3) Berg, J. v. d., Ooyen, A. v., Mantei, N., Schamböck, A., Grosveld, G., Flavell, R. A., and Weissmann, C.: Nature, **276**, 37 (1978).
4) Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H., and Honjo, T.: Proc. Natl. Acad. Sci. U.S.A., **76**, 4240 (1979).
5) Honjo, T., Obata, M., Yamawaki-Kataoka ,Y., Kataoka, T., Kawakami, Y., Takahashi, N., and Mano, Y.: Cell, **18**, 559 (1979).
6) Sakano, H., Rogers, J. H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R., and Tonegawa, S.: Nature (London), **277**, 627 (1979).
7) Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N., and Hood, L.: Proc. Natl. Acad. Sci. U.S.A., **76**, 857 (1979).
8) Tiemeier, D. C., Tilghman, S. M., Polsky, F. I., Seidman, J. G., Leder, A., Edgel, M. H., and Leder, P.: Cell, **14**, 237 (1978).
9) Crick, F.: Science, **204**, 264 (1979).
10) Robenberg, M., and Kramer, R. A.: Proc. Natl. Acad. Sci. U.S.A., **74**, 984 (1977).
11) Robertson, H. D., Dickson, E., and Dunn, J. J.: Ibid., **74**, 822 (1977).
12) Young, R. A., and Steiz, J. A.: Ibid., **75**, 3593 (1978).
13) Adetugbo, K.: J. Biol. Chem., **253**, 6068 (1978).
14) ——: J. Biol. Chem., **253**, 6076 (1978).
15) Birshtein, B. K., Preud'homme, J.-L., and Scharff, M. D.: Proc. Natl. Acad. Sci. U.S.A., **71**, 2478 (1974).
16) Preud'homme, J.-L., Birshtein, B. K., and Scharff, M. D.: Proc. Natl. Acad. Sci. U.S.A., **72**, 1427 (1975).
17) Francus, T., Charmgrongartana, B., Campbell, R., Scharff, M. D., and Birshtein, B. K.: J. Exp. Med., **147**, 1535 (1978).
18) Max, E. E., Seidman, J. G., and Leder, P.: Proc. Natl. Acad. Sci. U.S.A., **76**, 3450 (1979).
19) Sakano, S., Hüppi, K., Heinrich, G., and Tonegawa, S.: Nature (London), **280**, 288 (1979).
20) Seidman, J., Max, E. E., and Leder, P.: Nature (London), **280**, 370 (1979).
21) Swan, D., Aviv, H., and Leder, P.: Proc. Natl. Acad. Sci. U.S.A., **70**, 451 (1973).
22) Honjo, T., Packman, S., Swan, D., Nau, M., and Leder, P.: Ibid., **71**, 3659 (1974).
23) Maxam, A., and Gilbert, W.: Proc. Natl. Acad. Sci. U.S.A., **74**, 560 (1978).
24) Needleman, S. B., and Wunsch, C. D.: J. Mol. Biol., **48**, 443 (1970).

Gene, 9 (1980) 87—97

IMMUNOGLOBULIN γ1 HEAVY CHAIN GENE: STRUCTURAL GENE SEQUENCES CLONED IN A BACTERIAL PLASMID

(Mouse myeloma; dA-dT tailing; R-loop; lysine residue at C-terminus; 3′ untranslated sequence)

MASAHIRO OBATA, YURIKO YAMAWAKI-KATAOKA, NAOKI TAKAHASHI, TOHRU KATAOKA, AKIRA SHIMIZU, YOSHITAKE MANO, J.G. SEIDMAN*, B. MATIJA PETERLIN*, PHILIP LEDER* and TASUKU HONJO

*Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo, (Japan), and *Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institute of Health, Bethesda, MD 20205 (U.S.A.)*

(Received September 25th, 1979)
(Accepted October 8th, 1979)

SUMMARY

Immunoglobulin γ1 heavy chain cDNA was cloned into an *Escherichia coli* plasmid pCR1 and its nucleotide sequence was determined. The hybrid plasmid contained approx. 900-bases-long γ1 chain cDNA sequence, including the complete sequence of the C_{H2} and C_{H3} domains and the 3′ untranslated region, and a partial sequence of the C_{H1} domain. The nucleotide sequence predicts an extra lysine at the carboxyl-terminus of the γ1 chain. Comparison of the nucleotide sequence of 3′ untranslated regions of the immunoglobulin γ1 chain and κ light chain showed a significant homology although lengths are quite divergent.

INTRODUCTION

Cloning and characterization of immunoglobulin heavy chain cDNA are important not only to obtain a probe for cloning its chromosomal counterpart but also to obtain direct information which is not available from the analysis of the chromosomal gene. For example, the nucleotide sequence determination of cDNA clone defines the 3′ untranslated sequence and poly(A) additional site. The direct evidence for splicing of immunoglobulin mRNA precursors (Rabbits, 1978) is also obtained by comparison of nucleotide sequences between genomic DNA and cDNA.

We have constructed double-stranded immunoglobulin heavy chain cDNA which is complementary to γ1 chain mRNA purified from mouse myeloma

MOPC 31C (Ono et al., 1977), cloned the cDNA into a bacterial plasmid pCR1 and determined its nucleotide sequence. The cDNA clone includes the complete sequences of the C_{H2} and C_{H3} domains and 3′ untranşlated region, and a part of the C_{H1} domain. The cDNA clone has already been used for cloning the γ1 chain gene from the newborn mouse (Honjo, 1978; Honjo et al., 1979).

MATERIAL AND METHODS

(a) Materials

Reverse transcriptase from avian myelobastosis virus was provided by Dr. J.W. Beard. [γ-^{32}P] ATP (7000 Ci/mmol), [α-^{32}P] dCTP (3000 Ci/mmol) and ^{125}I were purchased from New England Nuclear (Boston, MA). Polynucleotide kinase and restriction endonucleases (*Hin*dIII, *Hin*fI, *Mbo*II, *Pst*I, *Sma*I and *Pvu*II) were purchased from New England Biolabs (Beverly, MA). *Eco*RI, *Hae*III, *Hap*II and *Hha*I, which were purified according to the method of Modrich and Zable (1976), Middleton et al. (1972), Takanami (1974) and Green et al., (1978), respectively. *Ava*II and *Acc*I were gifts from Dr. Nakanishi of Kyoto University and Dr. Kajiro of University of Tokyo, respectively. Terminal nucleotidyl transferase was provided by Dr. R.L. Ratliff.

(b) Construction of a hybrid plasmid

Preparation of double-stranded cDNA was carried out as described previously (Seidman et al., 1978). Starting with 10 μg of purified MOPC 31C mRNA (Ono et al., 1977), we obtained 1.6 μg of double-stranded cDNA, which was treated with S1 nuclease. Then poly(dA) stretches were attached to the double-stranded cDNA, and poly(dT) stretches were added to pCR1 DNA (Armstrong et al., 1977), which had been cleaved with *Eco*RI, by the action of terminal nucleotidyl transferase using the procedure described by Roychoudhury et al., (1976). The poly(dA)-tailed cDNA was annealed to the poly(dT)-tailed pCR1 to form circular hybrid molecules. The hybrid molecules were used to transform EK2 host χ1776. Finally, the transformants were screened for the immunoglobulin γ1 chain sequence using the procedure of Grunstein and Hogness (1975).

(c) In situ hybridization and hybridization probes

Transfer of DNA from agarose gel to a nitrocellulose filter was done according to Southern (1975). The filter was hybridized with ^{32}P-labeled cDNA as described (Kataoka et al., 1979). The cDNA complementary to the purified MOPC 31C mRNA was synthesized using [α-^{32}P] dCTP and reverse transcriptase (Honjo et al., 1974). The specific activity was 60 cpm/pg. Immunoglobulin heavy chain mRNA was purified as described previously (Ono et al., 1977). Iodination of the purified mRNA was done as described (Commerford, 1971). The specific activity of [^{125}I] mRNA was 20 cpm/pg.

(d) Electron microscopy

R-loop between pG1-6 DNA and MOPC 31C mRNA were formed with 100 μg/ml of DNA (digested with *Hin*dIII) and 50 μg/ml of mRNA in 70% formamide—0.1 M tricine pH 8.0—0.5 M NaCl—10 mM EDTA at 52°C for 16 h as described earlier (Tilghman et al., 1978). After 16 h SV40 DNA containing poly(dT) tail was added to a final concentration of 5 μg/ml and the mixture was incubated at 52°C for several hours before being spread and stained as described (Tilghman et al., 1978). Preparation of SV40 DNA-poly(dT) was described (Tilghman et al., 1978).

(e) DNA sequence analysis

pG1-6 DNA was digested with *Hha*I and the 1.8 kb long fragment was isolated by electrophoresis in 5% polyacrylamide gel. The *Hha*I fragment was digested with various restriction enzymes shown in Fig. 4, treated with bacterial alkaline phosphatase and labeled at the 5′ end with polynucleotide kinase in the presence of [γ-^{32}p] ATP (Maxam and Gilbert, 1977). After cleavage with a second restriction enzyme, fragments labeled at one end were isolated by polyacrylamide gel electrophoresis. Sequencing was performed according to the method of Maxam and Gilbert (1977). Four base-specific reactions were used (G, A>C, C+T, C) and samples were separated by a thin (0.5 mm) polyacrylamide gel (Sanger and Coulson, 1978).

RESULTS

Cloning of γ1 chain cDNA

The construction of double-stranded cDNA and its cloning were carried out as described previously (Seidman et al., 1978). We started with γ1 chain mRNA highly purified from MOPC 31C myeloma (Ono et al., 1977). The double-stranded DNA was formed by reverse transcriptase, *E. coli* DNA polymerase I and S1 nuclease, and joined to a plasmid vector pCR1 (*Eco*RI digested) using the poly(dA-dT) tailing procedure.

Using 0.2 μg of hybrid molecules to transform *E. coli* strain χ1776, we obtained 215 transformant clones. These 215 clones were grown on nitrocellulose filters, and hybridized to purified ^{125}I-labeled γ1 chain mRNA using the procedure of Grunstein and Hogness (1975). 16 of the 215 clones hybridized to ^{125}I-labeled mRNA. The DNA from these 16 clones were screened for their ability to protect γ1 chain cDNA (1800 bases long) from S1-nuclease digestion. We selected a clone whose DNA protected about 50% of the cDNA, and designated it as pCR1·G1-6 (abbreviated pG1-6).

Restriction enzyme digestion of cDNA clones

DNA of pCR1 and pG1-6 was digested with restriction endonucleases *Sma*I, *Pst*I or *Hha*I. Southern blots of agarose gels of these digests were hybridized with γ1 chain [^{32}P] cDNA. As shown in Fig. 1 the largest fragment (7.3 kb) of *Sma*I-digested pCR1 DNA disappeared in *Sma*I digests of pG1-6

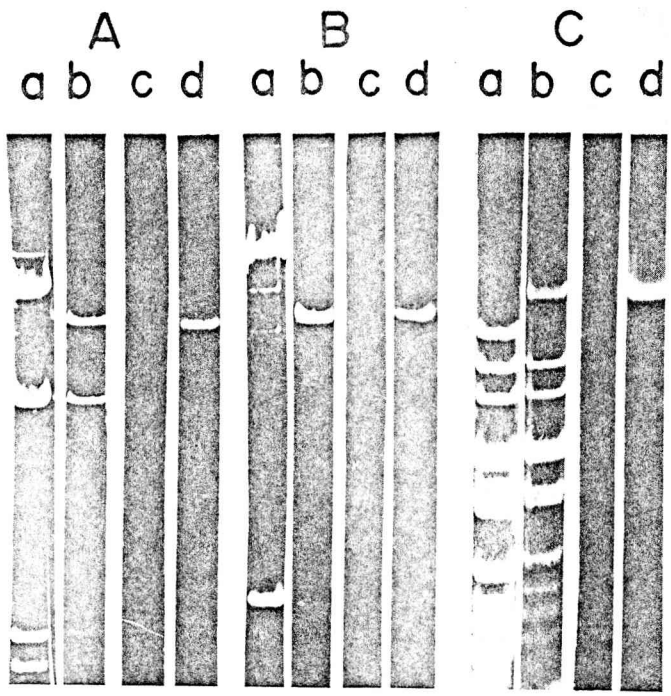


Fig. 1. In situ hybridization of restriction DNA fragments of pCR1·G1-6. DNAs of pCR1 and pCR1·G1-6 were digested with *Sma*I (A), *Pst*I (B), or *Hha*I (C) and electrophoresed in 1% agarose gels. The gels were stained with ethidium bromide and DNA was transferred to nitrocellulose filters according to Southern (1975). The filters were hybridized with MOPC 31C [^{32}P]cDNA. (a) and (b) are ethidium bromide-stained agarose gels of pCR1 and pCR1·G1-6, respectively. (c) and (d) are autoradiograms of Southern blots of pCR1 and pCR1·G1-6, respectively. *Sma*I digestion of pCR1 yields four fragments of 7.3, 2.9, 1.1 and 1.0 kb and *Pst*I cuts pCR1 into two fragments of 11 and 1.4 kb (Armstrong et al., 1977).

DNA and 2 fragments of about 6.0 and 1.9 kb in length appeared. These fragments were shown to contain the γ1 chain gene sequence by in situ hybridization although the 1.9 kb band was seen only by long exposure. Similarly, the largest fragment (11 kb) of *Pst*I-digested pCR1 was not seen in *Pst*I-digested pG1-6 whereas two fragments of about 6.1 and 6.0 kb in length were seen in the hybrid clone. The larger of the new fragments hybridized with γ1 chain [^{32}P]cDNA. The results indicate that pG1-6 contains the γ1 chain cDNA insert of approximately 1.1 to 0.6 kb in length and that the cDNA insert is cleaved by both *Sma*I and *Pst*I. The smaller of the new *Pst*I fragment may have a short cDNA sequence which does not show up in autoradiogram. *Hha*I digestion of pG1-6 showed the absence of the fragments corresponding to the first and seventh largest fragments of *Hha*I-digested pCR1 and the presence of a new fragment (about 1.8 kb) which hybridized with γ1 chain ^{32}P-cDNA. Since the largest fragment of *Hha*I digest contains a *Eco*RI site, deletion of the seventh largest fragment may be due to exonu-

cleolytic cleavage during dT tailing by terminal transferase. Four out of five clones analyzed lost the same *Hha*I-fragment.

R-loop study

R-loop was formed between MOPC 31C mRNA and pG1-6 DNA cleaved with *Hin*dIII (Fig. 2). Poly(dT) tailed SV 40 DNA, which was expected to hybridize to poly(A) sequence of mRNA, was added to show the direction of the inserted sequence. A series of R-loops was measured with the pCR1 as a duplex standard. The mean length (±S.D.) of R-loops in 27 molecules was calculated to be 760 ± 120 base pairs. The number is consistent with the results of restriction fragment analyses (Fig. 1). The direction of translation is from the left to the right in Fig. 2 and the *Hin*dIII site of pCR1 is closer to the 3′ end of the insert.

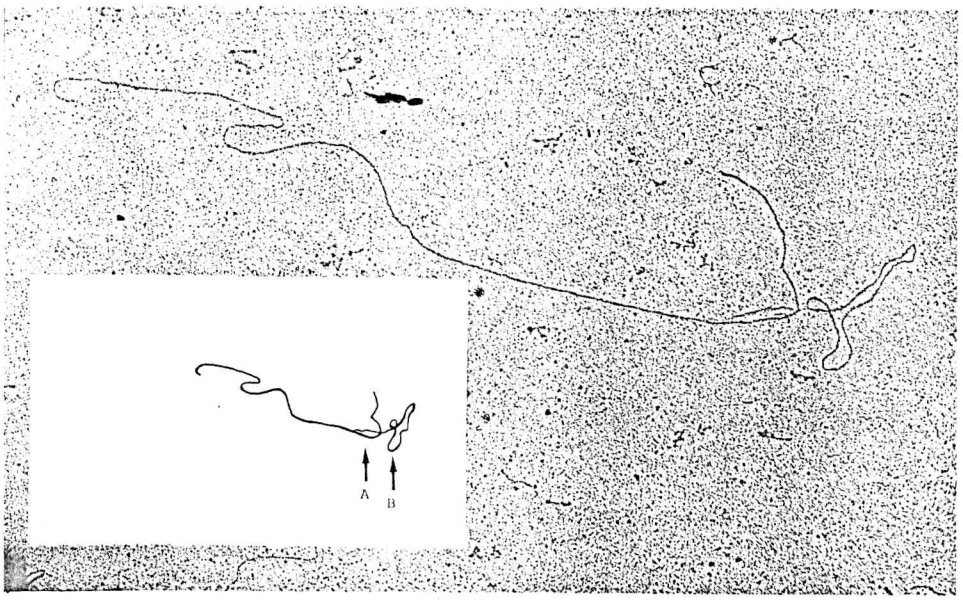


Fig. 2. Electron micrograph of an R loop between pCR1·G1-6 DNA and MOPC 31C mRNA. pCR1·G1-6 DNA cleaved with *Hin*dIII and MOPC 31C mRNA were hybridized to form an R loop (arrow A) under the conditions as described in METHODS. dT-tailed SV40 DNA was added as reference, which hybridized to poly(A) tail of mRNA as indicated by arrow B.

Strategy of nucleotide sequence determination

In order to sequence γ1 chain cDNA we isolated the 1.8-kb long *Hha*I fragment of pG1-6 and constructed a detailed restriction map as shown in Fig. 3. The determination of restriction sites was accomplished by digestion with several combinations of enzymes. The restriction sites used for terminal labeling were indicated by vertical bars and the ranges of sequences determined are shown by horizontal arrows at the bottom of Fig. 3.

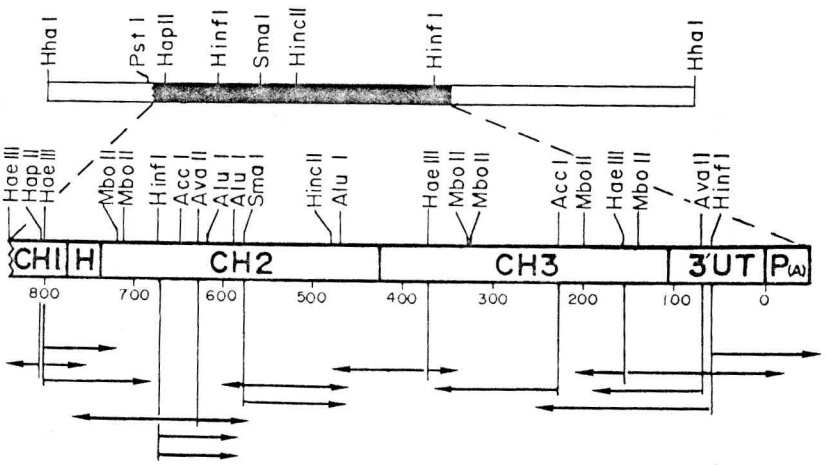


Fig. 3. Restriction map of the *Hha*I fragment of pCR1·G1-6. The top rectangle is a restriction map of the 1.8 kb *Hha*I fragment. A filled section indicates the inserted cDNA sequenced. The lower rectangle is a detailed map of the cDNA sequenced. Horizontal arrows at the bottom show the direction of sequencing and portion of read-off nucleotides. C_{H_1}, C_{H_2} and C_{H_3}, domains of $\gamma 1$ chain defined by amino acid sequence (Adetugbo, 1978); H, hinge region; 3′ UT, 3′ untranslated region; $P_{(A)}$, poly(A). Number is length in base-pairs beginning at the poly(A) additional site.

```
                                                  S  P                                   A  P
                                                  P  S  E  T  V  T  C  N  V  A  H  P  A  S
                                                 CCCAGCGAGACCGTCACCTGCAACGTTGCCCACCCGGCCAGCA
                                            843                                              801

 S  T  K  V  D  K  K  I  V  P  R  D  C  G  C  K  P  C  I  C  T  V  P  E  V  S  S  V  F  I  F  P  P  K
 GCACCAAGGTGGACAAGAAAATTGTGCCCAGGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCCCAAA
800                                                                                          701
          T     L        V
   P  K  D  V  L  T  I  T  L  T  P  K  V  T  C  V  V  V  D  I  S  K  D  D  P  E  V  Q  F  S  W  F  V
  GCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTA
700                                                                                          601
    N                                                 B           V           A
   D  D  V  E  V  H  T  A  Q  T  Q  P  R  E  E  Q  F  N  S  T  F  R  S  V  S  E  L  P  I  M  H  Q  D
  GATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACTTCCCATCATGCACCAGGACT
600                                                                                          501
                                                                                    K     R
  W  L  N  G  K  E  F  K  C  R  V  N  S  A  A  F  P  A  P  I  E  K  T  I  S  K  T  K  G  R  P  K  A  P
  GGCTCAATGGCAAGGAGTTCAAATGCAGGGTCAACAGTGCAGCTTTCCCTGCCCCCATCGAGAAAACCATCTCCAAAACCAAAGGCAGACCGAAGGCTCC
500                                                                                          401

   Q  V  Y  T  I  P  P  P  K  E  Q  M  A  K  D  K  V  S  L  T  C  M  I  T  D  F  F  P  E  D  I  T  V
  ACAGGTGTACACCATTCCACCTCCCAAGGAGCAGATGGCCAAGGATAAAGTCAGTCTGACCTGCATGATAACAGACTTCTTCCCTGAAGACATTACTGTG
400                                                                                          301
          S  D        A  P                                   D     D
   E  W  Q  W  N  G  Q  P  A  E  N  Y  K  N  T  Q  P  I  M  N  T  N  G  S  Y  F  V  Y  S  K  L  N  V
  GAGTGGCAGTGGAATGGGCAGCCAGCGGAGAACTACAAGAACACTCAGCCCATCATGAACACGAATGGCTCTTACTTCGTCTACAGCAAGCTCAATGTGC
300                                                                                          201

  Q  K  S  N  W  E  A  G  N  T  F  T  C  S  V  L  H  E  G  L  H  N  H  H  T  E  K  S  L  S  H  S  P  G
  AGAAGAGCAACTGGCAGGCAGGAAATACTTTCACCTGCTCTGTGTTACATGAGGCCCTGCACAACCACCATACTGAGAAGAGCCTCTCCCACTCTCCTGG
200                                                                                          101

    K  Ter
  TAAATGATCCCAGTGTCCTTGGAGCCCTCTGGTCCTACAGGACTCTGACACCTACCTCCACCCCTCCCTGTATAAATAAAGCACCCAGCACTGCCTTGGG
100                                                                                            1

  AAAAAAAAAA·····
```

Fig. 4. Nucleotide sequence of pCR1·G1-6. Nucleotide sequence is displayed 5′ to 3′ at the bottom line and numbered with 1 being the poly(A) additional site. Amino acids predicted from the nucleotide sequence are shown at the middle line by one letter code: A, alanine; B, aspartic acid or asparagine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine. Amino acid sequences of MOPC 21 (Adetugbo, 1978) are presented at the top line only where they disagree with the predicted residue.

Nucleotide sequence of γ1 chain structural gene

The nucleotide sequence determined is shown in Fig. 4. The sequence was correlated with the amino acid sequence of the γ1 chain of MOPC 21 (Adetugbo, 1978). The nucleotide sequence determined begins at the Pro codon of the 193rd residue which is close to the end of the C_{H1} domain and terminates at 93 nucleotides beyond the termination codon. The number of nucleotides determined is 843. The nucleotide sequence disagrees with the amino acid sequences of MOPC 21 at 18 residues out of 259 residues, which is shown at the top line in Fig. 4. A most interesting discrepancy is the presence of an additional Lys codon (AAA) at the carboxyl-terminus of the γ1 chain. The same codon was found in the chromosomal gene of the γ1 chain (Honjo et al., 1979).

DISCUSSION

The nucleotide sequence of pG1-6 demonstrates unequivocally that the hybrid plasmid contains the γ1 chain gene structural sequence. The total length of the cDNA insert seems to be longer than the number of nucleotides determined (843 bases). Although pG1-6 has a *Pst*I site in the inserted cDNA sequence as shown in Fig. 1, there is no *Pst*I site in the 843-bases long cDNA sequenced. Since we found a *Pst*I site in the chromosomal γ1 chain gene at 56 bases upstream to the cDNA sequenced (Honjo et al., 1979), we presume the total cDNA insert is at least 56 bases longer than that sequenced so far. The second largest *Pst*I fragment of pCR1·G1-6 did not hybridize with ^{32}P-cDNA (Fig. 1) because the majority of the cDNA probe was not long enough to reach the *Pst*I site.

We found an unexpected Lys codon at the carboxyl terminus of the γ1 chain gene. Recently we found that the γ2b chain gene also has an extra Lys codon at the carboxyl terminus (Yamawaki-Kataoka, Y. et al. unpublished data). Several explanations are possible for this discrepancy; (1) difference between myelomas (2) amino acid sequencing error (3) removal of the Lys residue after synthesis. Since all the γ chain proteins so far sequenced have Gly at the C-terminus (reviewed by Fudenberg et al., 1978), the first and the second possibilities are less likely than the third. The newly synthesized γ1 chain protein has a Lys residue at the carboxyl terminus, which may be removed eventually either during secretion or during circulation in the serum. The proteolytic enzyme of such specificity has not been described. The putative Lys cleaving enzyme should be highly specific because there is another Gly-Lys in the C_{H2} domain. The biological significance of the "shortlived" Lys residue is not clear.

The nucleotide sequence of pG1-6 established the complete sequence of the 3′ untranslated region of the γ1 chain gene. The 3′ untranslated sequence of the γ1 chain is 93 bases long, which is much shorter than that (208 bases) of κ light chain (Hamlyn et al., 1978). The nucleotide sequence before the poly(A) additonal site of the γ1 chain gene is similar but not identical to

that of the κ light chain gene as shown in Fig. 5. Both sequences share significant homology including a hexanucleotide AATAAA common to many eukaryotic mRNAs (Proudfoot and Brownlee, 1976; Efstratiadis et al., 1977; McReynolds et al., 1978; Hamlyn et al., 1978; Nakanishi et al., 1979). The results suggest that the 3′ untranslated sequences of the γ1 and κ chain genes may be derived from a common ancestor. The long stretch (126 bases) of extra nucleotides found in the κ chain gene consists of the repetition of simple patterns of arrangements of nucleotides (Hamlyn et al., 1978). The di-, tri- and tetra-nucleotides of the single base and the triplet CCT are abundant. Such a simple feature of the structure suggests that this portion of the sequence may be added by partial duplication or unequal crossing-over. The 3′ untranslated sequence of the γ1 chain gene also contains 8 CCTs. Codon choices for the γ1 chain C region gene have been assigned (Table I). A part of data was taken from the nucleotide sequence of the chromosomal γ1 chain gene (Honjo et al., 1979) for the portion where the nucleotide sequence of pG1-6 is unavailable. All codons are used at least once with the exception of UUG (Leu), CUA (Leu), UCG (Ser), CGU (Arg) and CGA (Arg). Other codons which are rarely used are UUU (Phe), UUA (Leu), CUU (Leu), GCG (Ala), CGC (Arg), and AGA (Arg). Certain codons are strongly preferred for some amino acids. UUC (Phe), CUG (Leu), CAG (Gln), GAG (Glu) and AGG (Arg) are examples. Nonrandom use of codons has been noted in other eukaryotic mRNAs (Efstratiadis et al., 1977; McReynolds et al., 1978; Nakanishi et al., 1979). As noted previously (McReynolds et al., 1978; Nakanishi et al., 1979) codons with the doublet CG are relatively rare and the preference of G or C at the third position is seen. Out of 342 codons 102 codons (29.8%) terminate in G, 125 (36.5%) in C, 67 (19.6%) in U and 48 (14%) in A.

We did not sequence the 5′ end of the inserted cDNA because the 5′ end fragment of *Hap*II-digested pG1-6 migrated as a diffuse band of 300 ~ 350

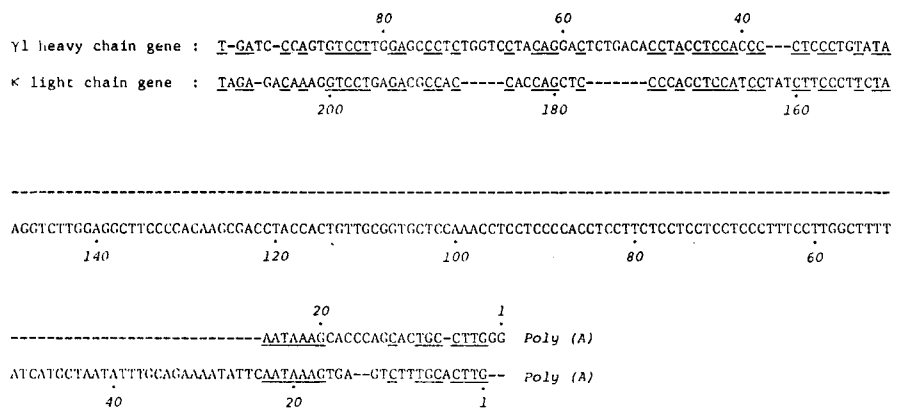


Fig. 5. Comparison of 3′ untranslated sequences between γ1 heavy chain and κ light chain genes. The two sequences are aligned to maximize homology by a modified computer program of Needleman and Wunsch (1970).

TABLE I

THE FREQUENCY OF USE OF EACH CODON FOR AMINO ACID OF γ1 HEAVY CHAIN CONSTANT REGION[a]

Amino acid	Codon	Number	Amino acid	Codon	Number	Amino acid	Codon	Number	Amino acid	Codon	Number
Ser	UCU	8	Thr	ACG	4	Leu	UUA	1	Gln	CAA	2
	UCC	6					UUG	0		CAG	11
	UCA	3	Lys	AAA	8		CUU	1			
	UCG	0		AAG	16		CUC	5	Ile	AUU	4
	AGU	3					CUA	0		AUC	7
	AGC	13	Arg	CGU	0		CUG	9		AUA	2
				CGC	1						
Val	GUU	2		CGA	0	Asn	AAU	5	Cys	UGU	4
	GUC	12		CGG	2		AAC	10		UGC	6
	GUA	4		AGA	1						
	GUG	14		AGG	17	Phe	UUU	1	His	CAU	2
							UUC	13		CAC	7
Pro	CCU	9	Ala	GCU	5						
	CCC	12		GCC	7	Gly	GGU	5	Tyr	UAU	2
	CCA	9		GCA	6		GGC	3		UAC	5
	CCG	2		GCG	1		GGA	4			
							GGG	1	Trp	UGG	6
Thr	ACU	10	Glu	GAA	3						
	ACC	12		GAG	14	Asp	GAU	7	Met	AUG	5
	ACA	5					GAC	6			

[a]A part of codons for C_{H1} is taken from the data of γ1 heavy chain genomic sequence (Honjo et al., 1979).

bases in length upon electrophoresis in 5% polyacrylamide gel. Digestion of the *Hap*II fragment with *Hae*III produced a 36 bases-long fragment and a shorter fragment of apparently heterogenous length, which, we presume, is due to the variable length of the dA-dT tail on this side. The diffuse *Hap*II band was obtained even when pG1-6 DNA was prepared from a single colony of bacterial cells carrying the plasmid pG1-6. It is likely that a long dA-dT tail serves as a target of the frequent recombination, resulting in the heterogeneous length of the dA-dT tail.

During preparation of this manuscript Rogers et al. (1979) reported the nucleotide sequence of γ1 chain cDNA which contains a part of the C_{H2} and C_{H3} domains. Their sequence is in agreement with ours except for bases 237, 243 and 505 where G is replaced by A.

ACKNOWLEDGEMENTS

We are grateful to Drs. M. Takanami and K. Sugimoto for instructing us in the nucleotide sequencing procedure, to Dr. Beard for kind supply of reverse transcriptase and to Dr. Miyata for allowing us to use his computer program. This investigation was supported in part by grants from Ministry of Education, Science and Culture of Japan.

REFERENCES

Adetugbo, K., Evolution of immunoglobulin subclasses: primary structure of a murine myeloma γ1 chain, J. Biol. Chem., 253 (1978) 6068—6075.

Armstrong, K.A., Hershfield, V. and Helinski, D.R., Gene cloning and containment properties of plasmid ColEl and its derivatives, Science 196 (1977) 172—174.

Commerford, S.L., Iodination of nucleic acids invitro, Biochemistry, 10 (1971) 1993—2000.

Efstratiadis, A., Kafatos, F.C. and Maniatis, T., The primary structure of rabbit β-globin mRNA as determined from cloned DNA, Cell, 10 (1977) 571—585.

Fudenberg, H.H., Pink, J.R.L., Wang, A.-C. and Douglas, S.D., Basic Immunogenetics, Oxford University Press, New York, 1978 p. 15.

Greene, P.J., Heynecker, H.L., Bolivar, F., Rodriguez, R.L., Betlach, M.C., Covarrubias, A.A., Backman, K., Russel, D.J., Tait, R. and Boyer, H.W., A general method for the purification of restriction enzymes, Nucl. Acids Res., 5 (1978) 2373—2380.

Grunstein, M. and Hogness, D.S., Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene, Proc. Natl. Acad. Sci. USA, 72 (1975) 3961—3965.

Hamlyn, P.H., Brownlee, G.G., Cheng, C.-C., Gait, M.J. and Milstein, C., Complete sequence of constant and 3′ noncoding regions of an immunoglobulin mRNA using the dideoxynucleotide method of RNA sequencing, Cell, 15 (1978) 1067—1075.

Honjo, T., Reorganization and expression of immunoglobulin genes in Y. Yamamura and T. Tada (Eds.), Immunology '78, Nakayamashoten, Tokyo, 1978, pp. 75—82.

Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, Tahahashi, N. and Mano, Y., Cloning and complete nucleotide sequence of mouse immunoglobulin γ1 chain gene, Cell, in press.

Honjo, T., Packman, S., Swan, D., Nau, M. and Leder, P., Organization of immunoglobulin genes: reiteration frequency of the mouse κ chain constant region genes, Proc. Natl. Acad. Sci. USA, 71 (1974) 3659—3663.

Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. and Honjo, T., Cloning immunoglobulin γ2b chain gene of mouse: characterization and partial sequencing, Proc. Natl. Acad. Sci. USA, in press.

Maxam, A.M. and Gilbert, W., A new method for sequencing DNA, Proc. Natl. Acad. Sci. USA, 74 (1977) 560—564.

McReynolds, L., O'Malley, B.W., Nisbet, A.D., Fothergill, J.E., Givol, D., Fields, S., Robertson, M. and Brownlee, G.G., Sequence of chicken ovalbumin mRNA, Nature, 273 (1978) 723—728.

Middleton, J.H., Edgell, M.H. and Hutchinson, C.A., Specific fragments of φX174 deoxyribonucleic acid produced by a restriction enzyme from *Haemophilus aegyptius* endonuclease Z[1], J. Virol., 10 (1972) 42—50.

Modrich, P. and Zable, D., *Eco*RI endonuclease: physical and catalytic properties of the homogeneous enzyme, J. Biol. Chem., 251 (1976) 5866—5874.

Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A.C.Y., Cohen, S.N. and Numa, S., Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipoprotein precursor, Nature, 278 (1979) 423—427.

Needleman, S.B. and Wunsch, C.D., A general method applicable to the search for similarities in the amino acid sequence of two proteins, J. Mol. Biol., 48 (1970) 443—453.

Ono, M., Kondo, T., Kawakami, M. and Honjo, T., Purification of immunoglobulin heavy chain messenger RNA by immunoprecipitation from the mouse myeloma tumor, MOPC-31C, J. Biochem. (Tokyo) 81 (1977) 949—954.

Proudfoot, N.J., Gillam, S., Smith, M. and Longley, J.I., Nucleotide sequence of the 3′ terminal third of rabbit α-globin messenger RNA; Comparison with human α-globin messenger RNA, Cell, 11 (1977) 807—818.

Rabbits, T.H., Evidence for splicing of interrupted immunoglobulin variable and constant region sequences in nuclear RNA, Nature 275 (1978) 291—296.
Rogers, J., Clarke, P. and Salser, W., Sequence analysis of cloned cDNA encoding part of an immunoglobulin heavy chain, Nucl. Acids Res., 6 (1979) 3305—3321.
Roychoudhury, R., Jay, E. and Wu, R., Terminal labeling and addition of homopolymer tracts to duplex DNA fragments by terminal deoxynucleotidyl transferase, Nucl. Acids Res., 3 (1976) 863—877.
Sanger, F. and Coulson, A.R., The use of thin acrylamide gels for DNA sequencing, FEBS lett., 87 (1978) 107—116.
Seeberg, P.H., Shine, J., Martial, J.A., Baxter, J.D. and Goodman, H.M., Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone, Nature, 270 (1977) 486—494.
Seidman, J.G., Edgell, M.H. and Leder, P., Immunoglobulin light-chain structural gene sequences cloned in a bacterial plasmid, Nature, 271 (1978) 582—585.
Shine, J., Seeburg, P.H., Martial, J.A., Baxter, J.D. and Goodman, H.M., Construction and analysis of recombinant DNA for human chorionic somatomammotropin, Nature, 270 (1977) 494—499.
Southern, E.M., Detection of specific sequences among DNA fragments separated by gel electrophoresis, J. Mol. Biol., 98 (1975) 503—517.
Takanami, M., Restriction endonucleases AP, GA, and H-1 from three *Haemophilus* strains, in Wickner, R.B. (Ed.), Methods in Molecular Biology, Vol. 7, Marcel Dekker, New York, 1974, pp. 113—133.
Tilghman, S.M., Tiemeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V., and Leder, P., Intervening sequence of DNA indentified in the structural portion of a mouse β-globin gene, Proc. Natl. Acad. Sci. USA, 75 (1978) 725—729.
Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M., Rat insulin genes: construction of plasmids containing the coding sequences, Science, 196 (1977) 1313—1319.

Communicated by K. Matsubara.

Cell, Vol. 18, 559-568, October 1979, Copyright © 1979 by MIT

Cloning and Complete Nucleotide Sequence of Mouse Immunoglobulin γ1 Chain Gene

Tasuku Honjo, Masanori Obata,
Yuriko Yamawaki-Kataoka, Tohru Kataoka,
Toshiaki Kawakami, Naoki Takahashi
and Yoshitake Mano
Department of Physiological Chemistry and Nutrition
Faculty of Medicine
University of Tokyo
Tokyo, Japan

Summary

The 6.6 kb DNA fragment coding for the immunoglobulin γ1 chain was cloned from newborn mouse DNA using λgt*WES*•λB as the EK2 vector. The complete nucleotide sequence (1823 bases) of the γ1 chain gene was determined. The cloned gene contained the entire constant region gene sequence as well as the poly(A) addition site, but not the variable region gene. The results indicate that the variable and constant region genes of immunoglobulin heavy chain are separated in newborn mouse DNA. The constant region genes of other gamma chains (that is, γ2a, γ2b and γ3) are not present in the cloned DNA fragment. The sequence demonstrates that the γ1 chain gene is interrupted by three intervening sequences at the junction of the domains and the hinge region, as previously shown in the γ2b and α chain genes and in the γ1 chain gene cloned from myeloma. The results suggest that the intervening sequence was introduced into the heavy chain gene before divergence of the heavy chain classes, and also support the hypothesis that the splicing mechanism has facilitated the evolution of eucaryotic genes by linking duplicated domains or prototype peptides not directly adjacent to one another. Comparison of the nucleotide sequence of the γ1 chain gene around the boundaries of the coding and intervening sequences with those of other mouse genes revealed extensive divergence, although short prevalent sequences of AG-GTCAG at the 5′ border of the intervening sequence and TCTGCAG-GC at the 3′ border were deduced. A limited homology of nucleotide sequences was found among domains and between the hinge region and the 5′ portion of the CH2 domain. Comparison of 3′ untranslated sequences from the γ1 and γ2b chain genes and the mouse major β-globin gene shows significant homology and a palindrome sequence surrounding the poly(A) addition site.

Introduction

An immunoglobulin polypeptide chain is composed of a variable (V) and a constant (C) region. The C portion of the immunoglobulin heavy chain contains three regions of homology, each of which is similar in size and homologous to the C region of the light chain. The homology regions seem to have a similar three-dimensional structure consisting of a compact domain which contributes to at least one active site (Edelman et al., 1969). Comparison of the amino acid sequences of each domain and light chain C region indicates that immunoglobulin heavy chains have evolved by gene duplication of a domain which shared a common ancestor with the C region of the light chain. Such analysis also suggests that gamma chain subclasses have evolved by gene duplication after the establishment of a three-domain structure (Grey, 1969; Milstein et al., 1974). Homologies between gamma chain genes were assessed by cross-hybridization using cDNA complementary to purified gamma chain mRNA (Yamawaki-Kataoka et al., 1979). The results show that gamma chain subclass genes share various extents of homology ranging from 20–60%, although gamma chain cDNAs do not have any significant homology with alpha and mu chain mRNAs.

We have recently cloned two immunoglobulin C genes of the gamma subclass, γ1 and γ2b from the newborn mouse (Honjo, 1978; Kataoka, Yamawaki-Kataoka and Honjo, 1979). Partial base sequence determination of the γ2b chain gene demonstrated that an intervening sequence (IVS) splits the heavy chain gene at the junction of the domains and the hinge region (Kataoka et al., 1979). Similar results were also reported for γ1 and α chain genes cloned from mouse myelomas (Early et al., 1979; Sakano et al., 1979). These results suggest that an IVS may have interrupted the immunoglobulin heavy chain gene at an early stage of its evolution, as has been proposed by Gilbert (1978) and also by Darnell (1978).

To trace the evolutionary history of immunoglobulin heavy chain genes, we have set out to determine the complete nucleotide sequences of the γ1 and γ2b chain genes and to compare them to each other. In this paper we report the cloning and complete nucleotide sequence of the γ1 chain C gene of the newborn mouse. The sequence not only allows us to identify the gene and precisely locate the three IVS that interrupt it, but also to characterize several interesting features of the coding region, the boundaries of the IVS and coding sequence and the 3′ untranslated sequence. A preliminary report has been published (Honjo, 1978).

Results and Discussion

Partial Purification of the Immunoglobulin γ1 Chain Gene

To enrich immunoglobulin gene sequences we used sequential purification steps of RPC-5 column chro-

matography and agarose gel electrophoresis. Approximately 40 mg of Eco RI-digested DNA derived from whole newborn mice were fractionated by RPC-5 chromatography. Aliquots of pooled fractions were electrophoresed in an agarose gel, and the Southern blot of the gel was hybridized with MOPC 31C (γ1) cDNA and autoradiographed. A marked 6.6 kb band was visible in fractions 5–7, as shown in Figure 1. No other bands were observed in any other fractions. Approximately 3.5 mg of DNA obtained in fractions 5–7 were applied to a preparative agarose gel. Aliquots of DNA fractions were again electrophoresed in a 1% agarose gel and blotted to a Millipore filter (Southern, 1975). The filter was hybridized with MOPC 31C (γ1) cDNA cloned in pCR1. As shown in Figure 2, the 6.6 kb fragment was found in fractions 14 and 15, in which the γ2b chain gene was simultaneously enriched (Kataoka et al., 1979). Faint 6.6 kb bands were observed in fractions 16–18, although ethidium bromide-stained bands of DNA were larger than 6.6 kb. Fraction 14 was concentrated by ethanol precipitation, and approximately 0.1 mg of DNA was obtained. The γ1 gene was estimated to have been purified about 100 fold from the total digest. This estimate is made from the relative proportion of the DNA in that pool after adjusting for the recovery of DNA and immunoglobulin-hybridizing sequences. Since Eco RI digestion of the mouse genome would be expected to generate about 6.7×10^5 fragments, fraction 14 is expected to contain one γ1 chain gene fragment among 6.7×10^3 fragments.

Isolation of a Recombinant Phage Containing the Immunoglobulin γ1 Chain C Gene

Purified outer fragments of λgt*WES*·λB were annealed with the DNA from fraction 14, ligated and used for transfection. Screening of clones was carried out according to the method of Benton and Davis (1977). After screening about 10^4 phages we found one recombinant containing the γ1 gene sequence, which was referred to as λgt*WES*·IgH2 (abbreviated IgH2). The recombinant phage was grown in large quantities and the DNA was extracted. DNA was hybridized with γ1 cDNA in solution, as shown in Table 1. γ1 cDNA hybridized with λgt*WES*·IgH2 but not with λgt*WES*·λB.

λgt*WES*·IgH2 DNA hybridized with MOPC 31C cDNA to an extent (73%) similar to that obtained for MOPC 70A mRNA containing the same C sequence and a different V sequence. The extent of hybridization is slightly greater than expected since the γ1 cDNA used and the C sequence including the 3′ untranslated sequence are about 1800 and 1100 bases in length, respectively (Honjo and Kataoka, 1978; see below). The extent of hybridization did not increase even when the amount of λgt*WES*·IgH2 DNA was increased 10 fold. The plasmid pG1-6, used as a control, contained about 900 bases of γ1 cDNA which hybridized to 50% of the cDNA, although the parental plasmid did not hybridize with the γ1 cDNA at all. The results suggest the λgt*WES*·IgH2 contains the C gene of the γ1 chain but does not encompass the V gene of the MOPC 31C heavy chain. To test whether λgt*WES*·IgH2 encodes

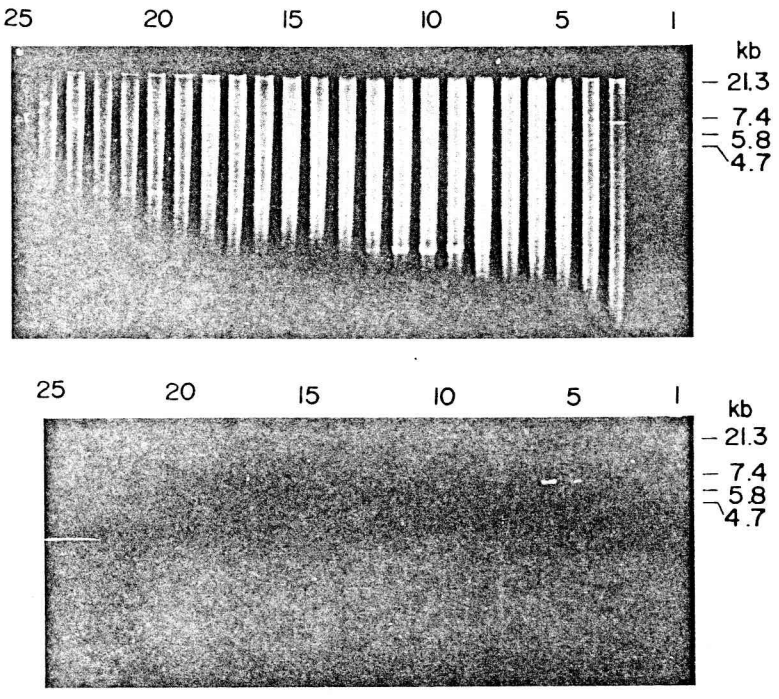


Figure 1. Fractionation of Eco RI-Digested Mouse DNA on RPC-5 Column Chromatography

(Upper section) An ethidium bromide-stained 1% analytical agarose gel of each pool. Migration is from top to bottom. (Lower section) An autoradiogram of the Southern blot which was hybridized with MOPC 31C ^{32}P-cDNA (γ1 chain). Eco RI-cleaved λCI857 was used as a size marker.

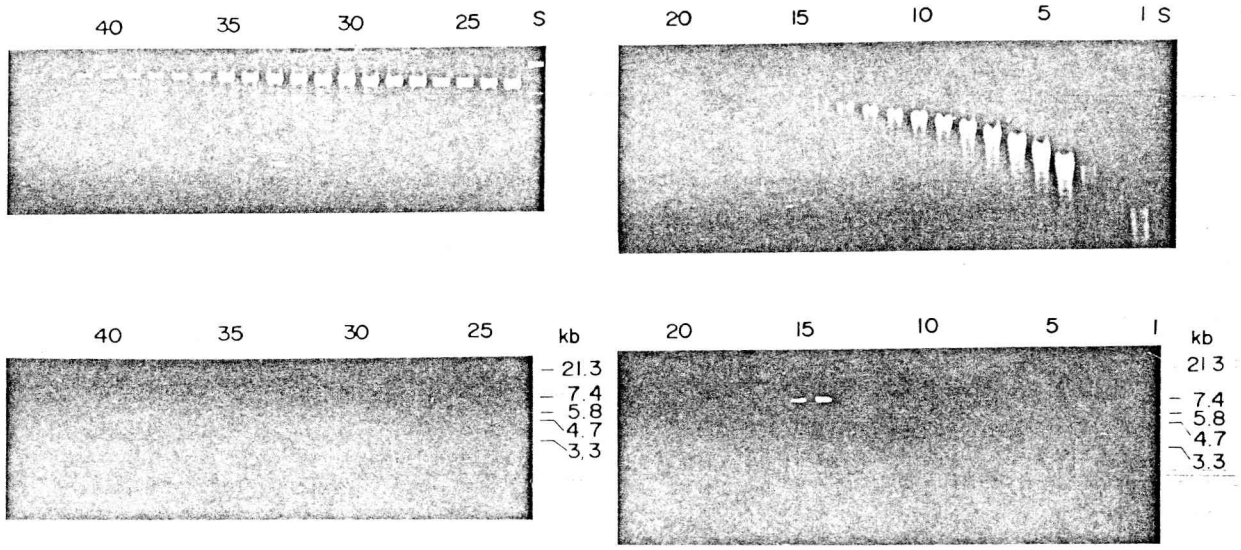


Figure 2. Fractionation of Eco RI-Digested Mouse DNA by Preparative Agarose Gel Electrophoresis

(Upper sections) An ethidium bromide-stained 1% analytical agarose gel of each fraction. Migration is from top to bottom. (Lower sections) An autoradiogram of the Southern blot which was hybridized with the ^{32}P-labeled Hha I fragment of pG1-6 (nick-translated). (S) Eco RI-cleaved λCI857 as size marker.

Table 1. Hybridization of λgt*WES*·IgH2 with MOPC 31C cDNA

DNA or mRNA	MOPC 31C cDNA % Hybridization
λgt*WES*·IgH2	73
λgt*WES*·λB	2
pG1-6	50
pCR1	2
MOPC 31C mRNA	100
MOPC 70A mRNA	79

Hybridization of DNA (1 μg) or mRNA (5 μg) was performed in the presence of 1300 cpm of MOPC 31C cDNA (γ1 chain), which is 1800 bases long. mRNA used in this experiment was partially purified from MOPC 31C and MOPC 70A myelomas which are γ1 chain producers.

Table 2. Hybridization of λgt*WES*·IgH2 with Gamma Chain cDNAs

cDNA	% Hybridization
MOPC 31C (γ1)	77
HOPC 1 (γ2a)	13
MPC 11 (γ2b)	13
J 606 (γ3)	15
MOPC 41 (κ)	2

Hybridization of 0.5 μg of λgt*WES*·IgH2 DNA was performed in the presence of 2000 cpm of each ^{3}H-cDNA.

the C genes of the other gamma subclasses, λgt*WES*·IgH2 DNA was hybridized to the other gamma chain cDNAs. Table 2 shows that λgt*WES*·IgH2 hybridized specifically to the γ1 cDNA. Small amounts of hybridization with γ2a, γ2b and γ3 cDNAs seem to be attributable to partial homology between the gamma chain C gene sequences (Yamawaki-Kataoka et al., 1979).

Restriction Site Mapping

λgt*WES*·IgH2 was digested with restriction endonuclease Eco RI and the 6.6 kb insert was isolated by agarose gel electrophoresis. Both the 6.6 kb insert and λgt*WES*·IgH2 were digested with Sac I, Bam HI, Hind II and Sma I, and the digests were electrophoresed in a 1% agarose gel. DNA was transferred to a nitrocellulose filter for hybridization as shown in Figure 3. Sac I digested λgt*WES*·IgH2 DNA into 23, 15.2, 2.5 and 0.6 kb fragments while the 6.6 kb insert was cleaved by Sac I into four fragments of 2.5, 2.0, 1.5 and 0.6 kb. Since Sac I has no cleavage sites in the outer fragment of λgt*WES*, 2.5 and 0.6 kb fragments produced by the digestion of λgt*WES*·IgH2 DNA must be derived from the 6.6 kb insert, and 2.0 and 1.5 kb fragments derived from the 6.6 kb insert are located at the ends of the insert. Bam HI cleaved the insert into four fragments of 3.35, 1.5, 1.18 and 0.62 kb. Comparison with known cleavage sites of Bam HI in λgt*WES*·λB (Leder, Tiemeier and Enquist, 1977) places the 3.35 and 0.62 kb fragments at the right and left ends of the insert, respectively. Other fragments produced by Sac I or Bam HI were ordered by the double digestion with Sac I and Bam HI. The remaining restriction sites were determined by comparing the sizes of new insert fragments to known adjacent enzyme sites in λgt*WES* left and right outer fragments and by cleavage with a combination of several restriction endonucleases. Hybridization with the MOPC 31C cDNA clone shows that most of the structural sequence is present on the largest fragment

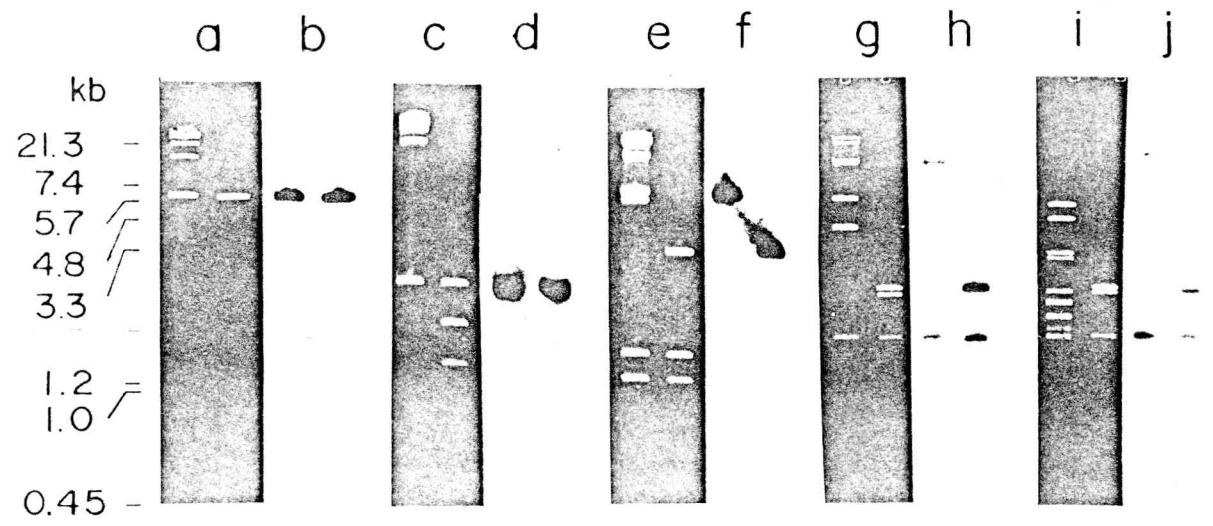


Figure 3. Digestion of λgt*WES*·IgH2 and Its Insert DNA with Various Restriction Endonucleases

Approximately 2 μg of λgt*WES*·IgH2 DNA and 0.36 μg of the insert DNA were cleaved and electrophoresed in a 1% agarose gel. The DNA was transferred to a Millipore filter (Southern, 1975) and hybridized to a nick-translated Hha fragment of pG1-6. Size markers used are fragments produced by Eco RI digestion of λCI857 DNA and Hae III digestion of Col E1 DNA. The following pairs of lanes are the ethidium bromide stain and autoradiogram of λgt*WES*·IgH2 and the insert DNA cleaved with Eco RI (a and b); Sac I (c and d); Bam HI (e and f); Sma I (g and h); Hind II (i and j).

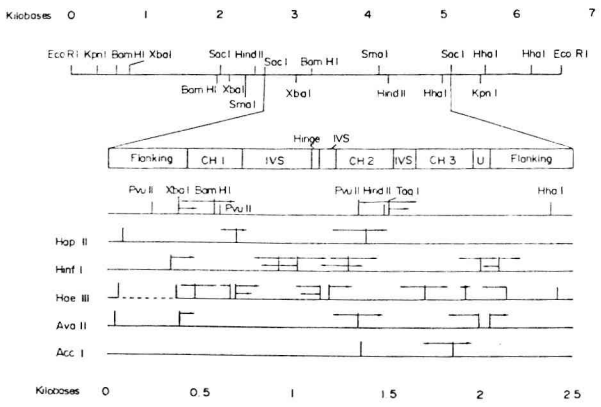


Figure 4. Diagram of Restriction Endonuclease Cleavage Sites in the γ1 Chain Gene and Strategy for Sequencing

The top line represents the cloned 6.6 kb Eco RI fragment in λgt*WES*·IgH2. The rectangle shows a magnified detail of the 2.5 kb Sac I fragment containing the entire structural gene of the γ1 chain. Cleavages used for sequencing are shown in the lower part of the figure. The horizonatal arrows indicate the direction and range of the sequence read. The fragments are aligned in the 5′ to 3′ direction relative to the coding sequence. (IVS) Intervening sequence; (U) untranslated sequence. The dashed line indicates the region where restriction sites are not determined.

of the Sac I digest, somewhere between 2.6 and 5.1 kb from the left end of the insert. The enzymes Sma I and Hind II cut within the coding sequence.

A more detailed restriction map was developed for the largest fragment (2500 bp) of Sac I digests containing the entire structural sequence. The mapping was carried out by combined cleavage with several restriction endonucleases and partial cleavage of the terminally ^{32}P-phosphorylated fragment (Smith and Birnstiel, 1976). Figure 4 summarizes the results obtained with this approach.

Sequencing Strategy

For DNA sequencing of IgH2, plasmid clone DNA (pBR322·IgH2) was digested simultaneously with Eco RI and Sac I. Since pBR322 does not have any restriction site for Sac I it was easy to isolate a 2.5 kb Sac I fragment on a 3% polyacrylamide gel. The 2.5 kb Sac I fragment was further digested with the restriction endonucleases indicated in Figure 4 and labeled at the 5′ ends with γ-^{32}P-ATP and polynucleotide kinase. The labeled DNA was either cut with another restriction enzyme to obtain fragments labeled only at one end or subjected to strand separation.

DNA sequencing was performed using the partial chemical degradation method of Maxam and Gilbert (1977), except that the thin polyacrylamide gel system (0.5 mm) of Sanger and Coulson (1978) was used to increase resolution. Regions were sequenced in both 5′ → 3′ and 3′ → 5′ directions whenever feasible. Restriction sites used for sequencing and the range of sequences obtained are shown in Figure 4.

Nucleotide Sequence of the γ1 Chain Gene

The complete sequence of the mouse γ1 chain gene is shown in Figure 5. The sequence starts 98 nucleotides before the structural sequence which begins with Ala at residue 121 of the γ1 chain sequence (Adetugbo, 1978). Our sequence ends 175 nucleotides after the termination codon (TGA) and includes a complete 3′ untranslated sequence. The total number of nucleotides in the segment sequenced is 1823.

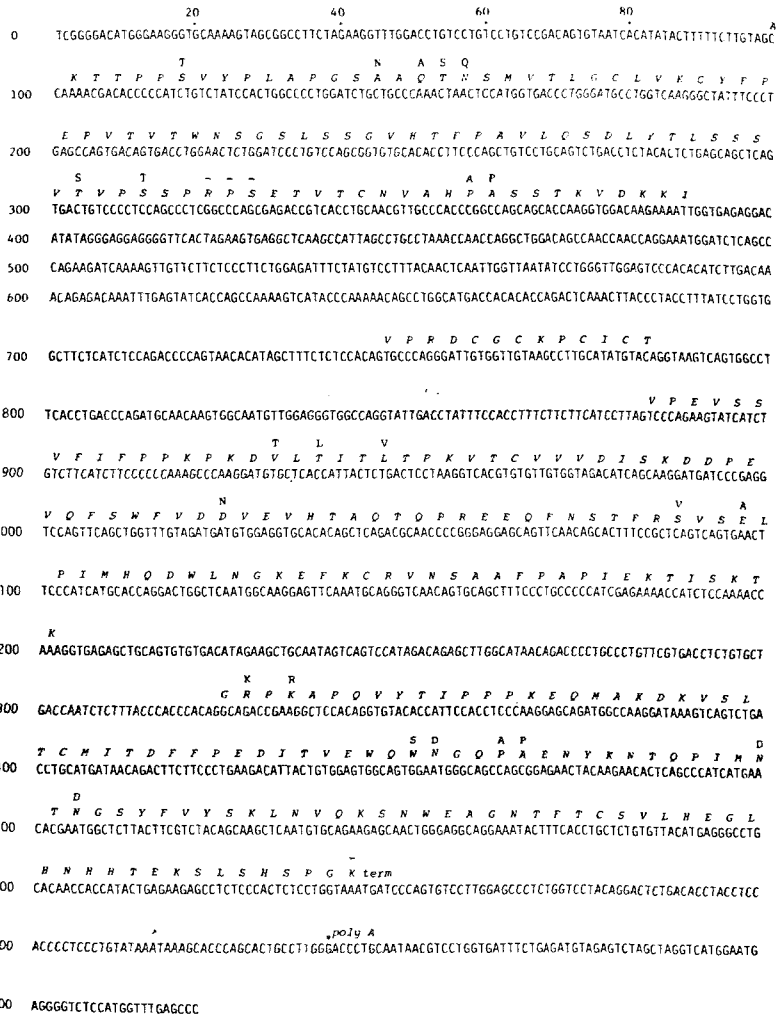


Figure 5. Nucleotide Sequence of the Mouse γ1 Chain Gene

The nucleotide sequence of the strand corresponding to the mRNA is displayed 5′ to 3′ on the bottom line. The amino acid predicted by the nucleotide sequence is shown above the coding sequence in italic letters. The amino acid sequence of MOPC 21 γ1 chain (Adetugbo, 1978) is presented on the top line. Only the amino acids which disagree with the predicted residue are shown. Dashes indicate deletions. (Term) refers to the termination codon UGA. [Poly(A)] indicates the site of poly(A) addition. Amino acids are expressed by a one letter code as follows: (A) alanine; (C) cysteine; (D) aspartic acid; (E) glutamic acid; (F) phenylalanine; (G) glycine; (H) histidine; (I) isoleucine; (K) lysine; (L) leucine; (M) methionine; (N) asparagine; (P) proline; (Q) glutamine; (R) arginine; (S) serine; (T) threonine; (V) valine; (W) tryptophan; (Y) tyrosine.

The nucleotide sequence can be correlated with the amino acid sequence already determined for the γ1 chain of MOPC 21 (Adetugbo, 1978). Our sequence disagrees at 27 of 324 residues. Nine of them are substitutions with other amino acids, including four Asn-Asp exchanges. Fourteen can be explained by displacement of amino acids. The nucleotide sequence predicts additional Arg-Pro-Ser residues between positions 194 and 195 of the MOPC 21 sequence. Since there are Trp-Pro-Ser residues at corresponding positions in the γ2a chain (Fougereau et al., 1976) and the codon for Trp (TGG) is similar to the codon for Arg (CGG) in our sequence, this discrepancy is likely to be an error of amino acid sequence. Another important discrepancy is a Lys residue predicted at the carboxy terminal end of the γ1 chain. We have sequenced the cDNA clone (pG1-6) synthesized from the γ1 chain mRNA of MOPC 31C, and the sequences for CH2 and CH3 domains, the hinge region and the 3′ untranslated sequence were identical to the genomic sequence (M. Obata and T. Honjo, unpublished data).

The nucleotide sequence clearly demonstrates that the V gene is separated from the C gene in newborn mouse DNA. Hybridization experiments (Table 1) and R loop studies (H. Yamagishi and T. Honjo, unpublished data) indicate that the V gene of the MOPC 31C heavy chain is not present in IgH2. We have recently cloned an 8.3 kb Eco RI fragment of mouse myeloma DNA, the 3′ terminal portion (6.3 kb) of which is identical to IgH2 by restriction site analysis (T. Kataoka, unpublished data). We are now investigating whether this conversion is related to a genetic rearrangement designed to bring the V and C genes closer together (Honjo and Kataoka, 1978).

Homology between the Domains and the Hinge Region

A computer-assisted search revealed limited homology between domains, as already indicated by the amino acid sequence. The similarity index (van den Berg et al., 1978) can be used as a rough estimate of homology between two sequences. The index expresses complete and random match as 100 and 25, respectively. As shown in Table 3, the similarity indexes between domains are 39.5–42.9, while the

Table 3. Homology between Domains[a]

	CH1 — CH2	CH1 — CH3	CH2 — CH3	β-Globin[b] — CH1	β-Globin — CH2	β-Globin — CH3
Matched bases	151	147	161	117	123	127
Deletions (bases)	117	126	102	108	129	129
Corrected length[c] (bases)	369	372	375	315	342	339
Similarity index[c]	40.9	39.5	42.9	37.1	36	37.5

[a] Sequences are aligned to maximize homology using a modified computer program described by Needleman and Wunsch (1970).
[b] The larger coding segment (residues 31–104) of major β-globin of mouse (Konkel et al., 1978).
[c] Similarity index (van den Berg et al., 1978) is computed as 100 − (% deletion × a + % transition × b + % transversion × c). Since the factors a, b and c are taken as 1, the similarity index is equal to the percentage of matched bases. The percentage is calculated relative to the corrected length, which is obtained by adding the number of base pairs and the number of deletions of the aligned sequences.

similarity indexes between two unrelated sequences—a segment of β-globin structural sequence (amino acid residues 31–104) and γ1 chain domains—are 36–37.5. We think that the latter numbers are an overestimate because the proportion of deletion is much higher. The most interesting homology occurs between the hinge region and the 5′ end of the CH2 domain, although the amino acid sequences of these regions are quite divergent. As shown in Figure 6, the homology extends into the intervening sequence. The similarity index of the sequence shown is 47.6. The results may suggest that the hinge region was derived from the 5′ end of the CH2 domain.

Intervening Sequence

Comparison of the nucleotide sequences of IgH2 and γ1 chain cDNA (M. Obata and T. Honjo, unpublished data) clearly shows that the γ1 chain C gene of newborn mouse is interrupted by three IVS of DNA dividing the coding sequence at the domain junctions and the hinge region as shown in Figure 4. The first IVS splits the γ1 chain gene after the end of the CH1 domain at residue 217. The second IVS begins after the codon for residue 230, separating the hinge region from the CH2 domain. The third IVS interrupts the gene between the codons for residues 337 and 338, separating the CH2 from the CH3 domain. The first, second and third IVS are 356, 98 and 121 bp long, respectively. No IVS was found in the 3′ untranslated sequence. The locations of the three IVS in the γ1 chain gene are virtually identical to those in the γ2b chain gene of the newborn mouse (Kataoka et al., 1979) and the γ1 chain gene of mouse myeloma (Sakano et al., 1979), and probably similar to those in the α chain gene of mouse myeloma (Early et al., 1979). The most probable conclusion to be drawn from these results is that the IVS was introduced into the heavy chain gene before divergence of the heavy chain classes. The results are also consistent with, although they do not prove, the hypothesis proposed independently by Gilbert (1978) and Darnell (1978) that the IVS and its splicing played a significant role in the evolution of eucaryotic genes. Gene duplication does not necessarily produce a contiguous domain structure, since duplicated domains are often interspersed by surrounding sequence which prevents the expression of a new function. Thus introduction of a splicing mechanism may have facilitated the evolution of a new functional gene and interspersed surrounding sequences may have diverged into present day IVS.

The first IVS contains three initiation and eight termination codons. The translation of the first IVS would produce a polypeptide 42 amino acids long, starting at base 95 after the end of CH1. The second IVS has two initiation codons, both of which are in correct phase with translation of the CH2 domain. The second and third IVS contain two and seven termination codons, respectively. All the IVS have termination codons in phase with the preceding coding block.

Comparison of Nucleotide Sequences at the Boundary of the IVS and Coding Sequence

It is evident that IVS are present in a nuclear mRNA precursor and that subsequent processing eliminates IVS by splicing structural sequences (Smith and Lingrel, 1978; Kinniburgh, Mertz and Ross, 1978; Tilghman et al., 1978). One might expect stable dyad symmetries to form the structural basis of the precise splicing machinery. However, the complete nucleotide sequence of the mouse β-globin gene did not reveal such sequences (Konkel, Tilghman and Leder, 1978). We have also searched for the sequence complementary to the IVS boundaries elsewhere in the flanking, coding and intervening sequences. So far we have been unable to find a region that forms stable duplexes with the boundary sequences.

It is also possible that the nucleotide sequences around the joints between IVS and the structural sequences may serve as signals for splicing enzyme(s). The nucleotide sequences surrounding the joints of IVS and structural sequences of known mouse genes are listed for comparison in Table 4. A direct comparison of the sequences listed allows us to derive the prevalent sequences AG-GTGAG at the 5′ border and TCTGCAG-GC at the 3′ border. These sequences share tetranucleotides—AG-GT at the 5′ border and

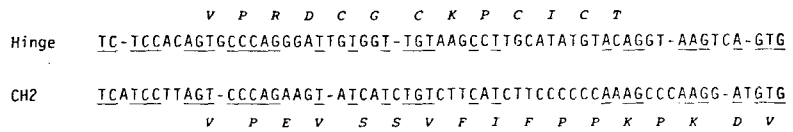


Figure 6. Comparison of the Hinge Region and 5′ Terminal Portion of the CH2 Domain

Two sequences with preceding intervening sequences are aligned to maximize homology. The common sequences are underlined. Amino acids encoded are shown by italic letters (see legend to Figure 5).

Table 4. Comparison of Nucleotide Sequences around Junctions between Coding and Intervening Sequences

Gene	5′ Junction		3′ Junction	Location
γ1 chain			TTTCTTGTAGCCAAAACGAC	VH:CH1
γ1 chain	AAGAAAATTGGTGAGAGGAC	.....	CTCTCCACAGTGCCCAGGGA	CH1:hinge
γ1 chain	ATATGTACAGGTAAGTCAGT	.....	TCATCCTTAGTCCCAGAATG	hinge:CH2
γ1 chain	AAAACCAAAGGTGAGCTGCA	.....	CCACCCACAGGCAGACCGAA	CH2:CH3
λ1 chain	CTCAGATCAGGTCAGCAGCC	.....	CTGTTTGCAGGGGCCATTTC	leader:VL
λ1 chain	ACTGTCCTAGGTGAGTCACT	.....	CATCCTGC-GGCCAGCCCAA	VL:CL
λ2 chain	CTCTGCTCAGGTCAGCAGCC	.....	CTGTTTGCAGGAGCCAGTTC	leader:VL
Major β-globin	CCCTGGGCAGGTTGGTATCC	.....	CCCTTTTTAGGCTGCTGGTT	position 30–31
Major β-globin	GAACTTCAGGGTGAGTCTGA	.....	ATTCCCACAGCTCCTGGGCA	position 104–105
Minor globin	TCAGGGTG			
α-globin	TC-AGGTAT			
Prevalent sequence	AGGTGAG		TCTGCAGGC	

Known mouse IVS border sequences are aligned to show maximum homology. The mouse λ chain sequences were determined by Tonegawa et al. (1978) and Bernard et al. (1978). The mouse globin sequences were reported by Konkel et al. (1978). The homologous nucleotides are underlined.

CAG-G at the 3′ border—with the proposed ancestral sequence (Breathnach et al., 1978; Konkel et al., 1978). The common sequences are dinucleotides, GT at the 5′ end and AG at the 3′ end of the IVS. Obviously such a short homology segment will not provide the strong specificity required for precise splicing.

Although we cannot determine exact splicing points for the γ1 chain gene, Breathnach et al. (1978) found that it was possible to define unique common excision-ligation points by aligning sequences of the joint region. They found GT at the 5′ end and AG at the 3′ end of the intervening sequences to be excised. This rule applies to all the boundary sequences of mouse genes thus far published, except the λ1 chain gene determined by Bernard, Hozumi and Tonegawa (1978). It should also be noted that the 5′ ends of IVS are rich in purines, especially G, and that the 3′ ends are rich in pyrimidines and rare in G. The base composition of the first and last 20 nucleotides of the IVS in the γ1 chain gene is shown in Table 5. A similar observation was made by Breathnach et al. (1978) and Konkel et al. (1978).

Table 5. Base Composition of the Boundary of the Coding and Intervening Sequences of the γ1 Chain Gene

Location of IVS	5′ End (%)				3′ End (%)			
	G	A	T	C	G	A	T	C
VH:CH1					5	25	50	20
CH1:hinge	45	40	10	5	5	20	30	40
Hinge:CH2	35	25	25	15	0	15	50	35
CH2:CH3	45	20	25	10	0	30	25	45

Base composition was calculated for 20 nucleotides at each end of the IVS excluding AG-GT.

Comparison of γ1 Chain Genes Cloned from Newborn Mouse and Myeloma

During the preparation of this manuscript, Sakano et al. (1979) published the sequences of the DNA surrounding the IVS boundaries of the γ1 chain gene cloned from mouse myeloma. Their sequence agrees with ours except for 9 bases scattered throughout the IVS. The most important discrepancy is the presence of GGT, a codon for Gly, at the 3′ end of the CH1 domain in our clone. Thus the splicing codon of their sequence is GAG(Val)-GTG(Val), another exception to GT-AG rule. There are several possible explanations for this discrepancy: differences between the undifferentiated gene of the newborn mouse and the differentiated gene of myeloma; mutations in the tumor cell; sequence error; or a combination of these factors. It would be interesting to see whether modification of the splicing frame is required for the expression of the γ1 chain gene in differentiated lymphocytes. Since the IVS is eventually excised, it seems less probable that replacement of several nucleotides in the IVS accompanies differentiation.

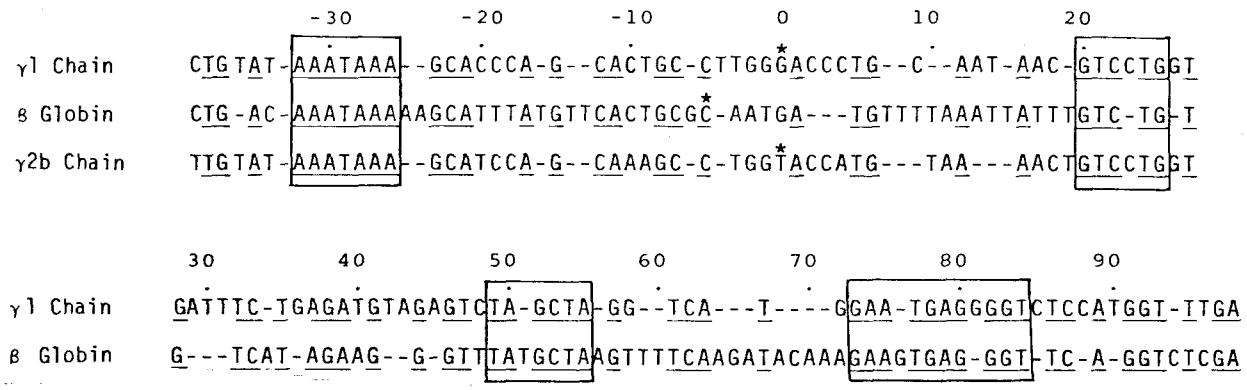


Figure 7. Comparison of the 3′ Terminal Portion of Mouse γ1 Chain, γ2b Chain and Major β-Globin Genes

Three sequences were aligned to maximize homology. Long homology regions are boxed. The common sequences are underlined. Dashes indicate artifical deletions required for maximal homology. The putative poly(A) addition site of the γ2b chain gene was assigned by comparison with the γ1 chain gene. (*) Poly(A) addition site.

γ1 Chain TGGGACCCT

γ2b Chain CCTGGT°ACCATG

β Globin TCACTGC°GCAATGA

Figure 8. Palindrome Sequences around Poly(A) Addition Sites

Palindrome sequences are underlined. (*) Poly(A) addition site; (O) center of symmetry.

3′ Flanking Sequence

Comparison of the nucleotide sequences of the γ1 chain gene and γ1 chain cDNA (pG1-6) indicates that the 3′ untranslated sequence is 93 nucleotides long, starting at TCC and ending at GGG (M. Obata and T. Honjo, unpublished data). We find GG at the poly(A) addition site instead of the GC found in many other eucaryotic mRNAs (Proudfoot and Brownlee, 1976; Hamlyn et al., 1978; McReynolds et al., 1978). The hexanucleotide AATAAA is found 20–25 bases before the poly(A) addition site, as is true of other eucaryotic mRNAs including globin, ovalbumin, immunoglobulin light chain and corticotropin–β–lipotropin (Proudfoot and Brownlee, 1976; Efstratiadis, Kafatos and Maniatis, 1977; Ullrich et al., 1977; Konkel et al., 1978; McReynolds et al., 1978; Nakanishi et al., 1979).

The nucelotide sequence around the poly(A) addition site in the γ1 chain gene was compared with that of the mouse major β-globin gene (Konkel et al., 1978). The similarity indexes of the 3′ untranslated sequences before and after the poly(A) addition site are 46 and 49, respectively, showing significant homology compared to the 36–37 index for the coding sequences of the γ1 chain and the β-globin genes (Table 3). As shown in Figure 7, with proper deletions there is significant homology between two sequences which have evolved independently for hundreds of millions of years. A partial sequence of the γ2b chain gene is also presented. Four homology regions are boxed. The first homology region is the hexanucleotide mentioned above. The second homology region at positions 20–25 is the most interesting because it has a symmetric sequence, GTC·CTG or GT(C)TG. There are AT-rich sequences between the poly(A) addition site and the second homology region as shown in the 5S RNA gene (Korn and Brown, 1978). [It is worth noting that there are palindrome sequences around the poly(A) addition sites of these genes, as shown in Figure 8.] It is not clear where RNA polymerase II terminates the transcription of these genes or how a primary transcript is processed, given poly(A) at the 3′ ends. The significant homology of the nucleotide sequences, AT-rich sequences and palindrome sequences around the poly(A) addition sites of two genes suggest that such sequences might be a signal for poly(A) addition or transcriptional termination or both.

Experimental Procedures

Materials

γ-^{32}P-ATP (spec. act. > 5500 Ci/mmole) and α-^{32}P-dCTP (spec. act. > 300 Ci/mmole) were purchased from New England Nuclear (Boston, Massachusetts). Restriction endonucleases (Kpn I, Xba I, Hinf I, Pvu II and Taq I) and polynucleotide kinase were from New England Biolabs (Beverly, Massachusetts).

Eco RI, Sac I, Hha I, Hap II and Bam HI were purified according to the methods of Yoshimori (1971), R. J. Roberts (unpublished data), Greene et al. (1978), Sugisaki and Takanami (1973) and Greene et al. (1978), respectively. T4 DNA ligase and Sma I were gifts from M. Takanami (Kyoto University). Ava II and Acc I were supplied by S. Nakanishi (Kyoto University) and Y. Kajiro (University of Tokyo), respectively. Bacterial alkaline phosphatase was obtained from Worthington Biochemicals and dialyzed against 0.5 M Tris–HCl (pH 8.0) containing 50% glycerol to remove ammonium sulfate. PRC-5 resin was a gift from P. Leder (NIH).

Partial Purification of the γ1 Chain Gene

Whole newborn mice (sacrificed within 24 hr after birth) were homogenized in 50 mM Tris·HCl (pH 8.0), 10 mM EDTA, 0.15 M NaCl with a Waring blender for 30 sec at 4°C. Sodium dodecylsulfate to 0.3%

and $NaClO_4$ to 0.5 M were added to a whole homogenate. The homogenate was extracted with 0.5 vol each of water-saturated phenol and a chloroform-isoamyl alcohol mixture (24:1). The water phase was separated by centrifugation and dialyzed against 0.1 × SSC (1 × SSC is 0.15 M NaCl, 0.015 M Na citrate) for 24 hr. RNAase was added to 100 μg/ml and digestion was continued for 24 hr in a dialysis bag. DNA was extracted with the phenol-chloroform mixture as above and dialyzed extensively against 0.1 × SSC. High molecular weight DNA was digested with Eco RI and fractionated on a PRC-5 column as described (Pearson, Weiss and Kelmers, 1971; Tilghman et al., 1977). DNA containing the immunoglobulin heavy chain gene was assayed by in situ hybridization as described below. Fractions which hybridized with the γ1 chain cDNA by in situ hybridization were pooled, concentrated by ethanol precipitation and further purified by preparative agarose gel electrophoresis as described previously (Polsky et al., 1978).

Preparation and Screening of Recombinant Phage

Bacteriophage λgt*WES*·λB (Leder et al., 1977) was used as the EK2 vector and propagated in ED8656. Cloning experiments were carried out in a P2 facility according to the proposed Japanese Guidelines (1978) and revised NIH Guidelines (1976). Ligation of λgt*WES* outer fragments and purified mouse DNA and transfection of recombinants were carried out as described (Tiemeier, Tilghman and Leder, 1977). Sets of about 200 plaques were transferred directly from LB broth agar plates to Millipore filters according to the method of Benton and Davis (1977). The Millipore filters were hybridized to a ^{32}P-labeled Hha I fragment of the γ1 cDNA cloned in pCR1 (pG1-6) and then autoradiographed. Recombinant phage DNA was prepared as described (Tiemeier et al., 1977). The 6.6 kb insert DNA was isolated by agarose gel electrophoresis and recloned into pBR322 by ligation (Bolivar et al., 1977). Plasmid DNA was purified according to the procedure of Clewell and Helinski (1969).

Hybridization

DNA fixed on a Millipore filter was hybridized to an appropriate probe as described by Jeffreys and Flavell (1977) except that 1 M NaCl, 50 mM Tris·HCl (pH 7.4), 10 mM EDTA replaced 6 × SSC. The probes used were ^{32}P-cDNA (spec. act. 62 cpm/pg) complementary to MOPC 31C mRNA (Ono et al., 1977) or a nick-translated ^{32}P-labeled Hha I fragment (spec. act. 20 cpm/pg) of pG1-6. A Cot value of 2×10^{-3} was reached with respect to the γ1 chain sequence. When ^{32}P-cDNA was used as a probe, oligo(dT)-cellulose bound mRNA from MOPC 511 myeloma (IgA producer) was added to 10 μg/ml. The plasmid pG1-6 is a hybrid of pCR1 (Armstrong, Hershfield and Helinski, 1977) and contains approximately 900 bases of γ1 chain cDNA sequence. The γ1 chain insert was isolated as the Hha I fragment of 1.8 kb in length. Detailed characterization of the plasmid will be presented elsewhere. Nick translation of DNA was carried out as described (Maniatis, Jeffreys and Kleid, 1975). Solution hybridization was carried out as described (Honjo et al., 1974; Kataoka et al., 1979).

Preparation of 5′ Terminally Labeled Restriction Fragments

DNA fragments were digested with restriction enzymes, treated with bacterial alkaline phosphatase and phosphorylated with T4 polynucleotide kinase as described (Maxam and Gilbert, 1977). After cleavage with a second restriction enzyme, fragments labeled at only one end were isolated by polyacrylamide gel electrophoresis. In some cases, fragments labeled at both ends were first separated by polyacrylamide gel electrophoresis, eluted and then either digested with a second restriction enzyme or subjected to strand separation. Electrophoresis was performed in 150 × 150 × 2 mm slab gels of 5–15% acrylamide (1/30 as bisacrylamide) in 50 mM Tris–borate (pH 8.3) and 2.5 mM EDTA. Samples to be eluted were cut out, put in 10 mM Tris·Cl (pH 8.0), 5 mM EDTA in a dialysis bag and electrophoresed for 4–16 hr at 120 mA. The eluted DNA was purified over Sephadex G100 and concentrated by evaporation.

DNA Sequence Analysis

Sequencing was performed according to the method of Maxam and Gilbert (1977) using 20% urea-acrylamide gels (3 hr at 1600 V) and 8% gels (5 and 3 hr at 1800 V). Thin gels (0.5 mm) were used according to the procedure of Sanger and Coulson (1978) to increase resolution. Four base-specific reactions were used (G, A > C, C + T, C), and samples were loaded in 90% formamide rather than urea–NaOH. The limit of the accuracy of the sequence determined by the chemical modification method was critically evaluated by Konkel et al. (1978), and we reached essentially similar conclusions.

Acknowledgments

We are grateful to Drs. Y. Suzuki and M. Obinata for their critical readings of the manuscript. We thank Dr. T. Miyata for computer-assisted calculations of sequence homology. This investigation was supported in part by grants from the Ministry of Education, Science and Culture of Japan.

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. Section 1734 solely to indicate this fact.

Received July 10, 1979

References

Adetugbo, K. (1978). J. Biol. Chem. *253*, 6068–6075.

Armstrong, K. A., Hershfield, V. and Helinski, D. R. (1977). Science *196*, 172–174.

Benton, W. D. and Davis, R. W. (1977). Science *196*, 180–182.

Bernard, O., Hozumi, N. and Tonegawa, S. (1978). Cell *15*, 1133–1144.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. and Falkow, S. (1977). Gene *2*, 95–113.

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978). Proc. Nat. Acad. Sci. USA *75*, 4853–4857.

Clewell, D. and Helinski, D. (1969). Proc. Nat. Acad. Sci. USA *62*, 1159–1166.

Darnell, J. E., Jr. (1978). Science *202*, 1257–1260.

Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. and Hood, L. (1979). Proc. Nat. Acad. Sci. USA *76*, 857–861.

Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, U. and Waxdal, M. J. (1969). Proc. Nat. Acad. Sci. USA *63*, 78–85.

Efstratiadis, A., Kafatos, F. C. and Maniatis, T. (1977). Cell *10*, 575–585.

Fougereau, M., Bourgois, A., de Pravel, C., Rocca-Serra, J. and Schiff, C. (1976). Ann. Immunol. (Inst. Past.) *127C*, 607–631.

Gilbert, W. (1978). Nature *271*, 501.

Greene, P. J., Heynecker, H. L., Bolivar, F., Rodriguez, R. L., Betlach, M. C., Covarrubias, A. A., Backman, K., Russel, D. J., Tait, R. and Boyer, H. W. (1978). Nucl. Acids Res. *5*, 2373–2380.

Grey, H. M. (1969). Adv. Immunol. *10*, 51–105.

Hamlyn, P. H., Brownlee, G. G., Cheng, C.-C., Gait, M. J. and Milstein, C. (1978). Cell *15*, 1067–1075.

Honjo, T. (1978). Immunology '78, Y. Yamamura and T. Tada, eds. (Tokyo: Nakayama-shoten), pp. 75–82.

Honjo, T. and Kataoka, K. (1978). Proc. Nat. Acad. Sci. USA *75*, 2140–2144.

Honjo, T., Packman, S., Swan, D. and Leder, P. (1974). Proc. Nat. Acad. Sci. USA *71*, 3659–3663.

Japanese Scientific Monthly (1978). *31*, 626–647.

Jeffreys, A. J. and Flavell, R. A. (1977). Cell *12*, 429–439.

Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. and Honjo, T. (1979). Proc. Nat. Acad. Sci. USA, in press.

Kinniburgh, A. J., Mertz, J. E. and Ross, J. (1978). Cell *14*, 681–693.

Konkel, D. A., Tilghman, S. M. and Leder, P. (1978). Cell *15*, 1125–1132.

Korn, L. J. and Brown, D. D. (1978). Cell *15*, 1145–1156.

Leder, P., Tiemeier, D. and Enquist, L. (1977). Science *196*, 175–177.

McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M. and Brownlee, G. G. (1978). Nature *273*, 723–728.

Maniatis, T., Jeffreys, A. and Kleid, D. G. (1975). Proc. Nat. Acad. Sci. USA *72*, 1184–1188.

Maxam, A. M. and Gilbert, W. (1977). Proc. Nat. Acad. Sci. USA *74*, 560–564.

Milstein, C., Adetugbo, K., Cowan, N. J. and Secher, D. S. (1974). Prog. Immunol. II, *1*, pp. 157–168.

Nakanishi, S., Inoue, A. Kita, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N. and Numa, S. (1979). Nature *278*, 423–427.

Needleman, S. B. and Wunsch, C. D. (1970). J. Mol. Biol. *48*, 443–453.

Ono, M., Kondo, T., Kawakami, M. and Honjo, T. (1977). J. Biochem. (Tokyo) *81*, 949–954.

Pearson, R. L., Weiss, J. F. and Kelmers, A. D. (1971). Biochim. Biophys. Acta *228*, 770–774.

Polsky, F., Edgell, M. H., Seidman, J. G. and Leder, P. (1978). Anal. Biochem. *87*, 397–410.

Proudfoot, N. J. and Brownlee, G. G. (1976). Nature *263*, 211–214.

Recombinant DNA Reserach Guidelines, Federal Register, September 9, 1976.

Sakano, H., Rogers, J. H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S. (1979). Nature *277*, 627–633.

Sanger, F. and Coulson, A. R. (1978). FEBS Letters *87*, 107–116.

Smith, H. O. and Birnstiel, M. L. (1976). Nucl. Acids Res. *3*, 2387–2398.

Smith, K. and Lingrel, J. B. (1978). Nucl. Acids Res. *5*, 3295–3301.

Southern, E. M. (1975). J. Mol. Biol. *98*, 503–517.

Sugisaki, H. and Takanami, M. (1973). Nature *246*, 138–140.

Tiemeier, D. C., Tilghman, S. M. and Leder, P. (1977). Gene *2*, 173–191.

Tilghman, S. M., Tiemeier, D. C., Polsky, F., Edgell, M. H., Seidman, J. G., Leder, A., Enquist, L. W., Norman, B. and Leder, P. (1977). Proc. Nat. Acad. Sci. USA *74*, 4406–4410.

Tilghman, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P. and Weissmann, C. (1978). Proc. Nat. Acad. Sci. USA *75*, 1309–1313.

Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. and Gilbert, W. (1978). Proc. Nat. Acad. Sci. USA *75*, 1485–1489.

Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutler, W. and Goodman, H. M. (1977). Science *196*, 1313–1319.

van den Berg, J., van Ooyen, A., Mantei, N., Shamböck, A., Grosveld, G., Flavell, R. A. and Weissmann, C. (1978). Nature *276*, 37–44.

Yamawaki-Kataoka, Y., Sato, K., Shimizu, A., Kataoka, T., Mano, Y., Ono, M., Kawakami, M. and Honjo, T. (1979). Biochemistry *18*, 490–494.

Yoshimori, R. N. (1971). Ph.D. thesis, University of California, San Francisco, California.

ORGANIZATION AND REORGANIZATION OF IMMUNOGLOBULIN HEAVY CHAIN GENES- ALLELIC DELETION MODEL

Tasuku Honjo and Tohru Kataoka
Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo

We have assessed the number of times the gene sequence encoding constant regions of immunoglobulin heavy chains, γ1 and γ3 are represented in mouse genome by hybridization kinetic analysis. Both γ1 and γ3 genes are present in one copy per haploid genome in normal tissues and myelomas producing IgM or IgG3. IgG1-producing myelomas, however, contain one copy of the γ1 gene and 0.5 copy of the γ3 gene per haploid genome. In myelomas producing IgG2b, IgG2a or IgA, both γ1 and γ3 genes are represented 0.5 time per haploid genome. In order to account for the results we propose the allelic deletion model: (A) The specific deletion of heavy chain constant region genes accompanies the recombination of a variable region gene to a constant region gene. (B) The portion of the chromosome which resides between two joining sequences is excised out of the chromosome. (C) The recombination occurs on one of the alleles. Based on this model we propose that heavy chain genes are arranged on one chromosome in the following order; variable region genes, unknown spacer sequence; μ, γ3, γ1, γ2b, γ2a, and α.

I INTRODUCTION

Organization of immunoglobulin genes has been a focus of recent investigations in eukaryote molecular biology as well as immunology. Purification and reverse transcription of immunoglobulin light (L) chain mRNA have provided a powerful probe for nucleic acid hybridization studies on the particular genes. Extensive kinetic studies (1-9) have produced general agreement on important points of the organization of light chain genes. The kappa L chain constant (C) region gene sequence is present in very few copies - probably no more than two - per haploid genome. Since a contiguous mRNA

ISBN 0-12-370650-5

molecule contains both variable (V) and C region sequences (10), the above results strongly support the V-C recombination argument proposed by Dreyer and Bennett (11).

In contrast there have been few reports of molecular studies on organization of heavy (H) chain genes (12,13). This is mainly because purification of H chain mRNA is much more difficult than that of L chain mRNA. Recently we have succeeded in purification of H chain mRNA from mouse myelomas of several classes by immunoprecipitation of polysomes (14). Using these probe we have measured the number of H chain gene sequences and shown deletion of H chain constant region genes in mouse myelomas which produce immunoglobulins of specific classes. Based on these results we will propose the allelic deletion model for the mechanism of V_H-C_H recombination.

Genetic studies have shown that C_H genes of six classes and subclases, i.e., μ, $\gamma 1$, $\gamma 2a$, $\gamma 2b$, $\gamma 3$ and α, are clustered in such a narrow segment of one chromosome as there have been found no recombinants within the C_H region among more than 2,000 hybrid mice (15). These C_H genes share a family of V_H genes which sits on the same chromosome and is separated probably from the C_H genes by spacer sequences.

II CHARACTERIZATION OF PROBES

H chain mRNAs were purified from MOPC 31C ($\gamma 1$) and J606 ($\gamma 3$) myelomas by immunoprecipitation of polysomes as described previously (14). The mRNA preparation obtained ran as an homogeneous band upon polyacrylamide gel electrophoresis in formamide with molecular weight of 700,000, i.e. 2,000 bases. In wheat germ cell-free system $\gamma 1$ and $\gamma 3$ mRNAs directed synthesis of proteins which are slightly smaller than the authentic H chain proteins containing polysaccharide (14). Tryptic peptides derived from the *in vitro* synthesized products are similar to those obtained from the authentic proteins.

The mRNAs were reverse-transcribed with RNA-dependent DNA polymerase derived from avian myeloblastosis virus. The purities of MOPC 31C and J606 mRNAs were also determined by hybridization kinetics with respective cDNA. The $C_rt1/2$ values obtained for MOPC 31C and J606 mRNAs were 3.4×10^{-4} and 3.3×10^{-4}, respectively, as compared to a $C_rt1/2$ value of 1.8×10^{-4} for α and β globin mRNA under the same conditions (14). Given the genetic complexities of 1,200 bases and 2,000 bases for globin and H chain mRNAs, respectively, comparison of the $C_rt1/2$ values indicates that the purities of MOPC 31C and J606 mRNAs are 97% and 99%, respectively.

MOPC 31C and J606 cDNAs were purified by hybridization to homologous mRNA and/or by alkaline sucrose density gradient. The cDNA preparations employed for the present study are 1,500 bases long as determined by polyacrylamide gel electrophoresis in formamide. Since we do not see any difference in the extent of hybridization of cDNA between homologous and heterologous H chain mRNAs of the identical class, the cDNA preparations represent most of the C gene sequences but not of the V gene sequences. Diagramatic representation of H chain mRNA and its cDNA is shown in Fig. 1.

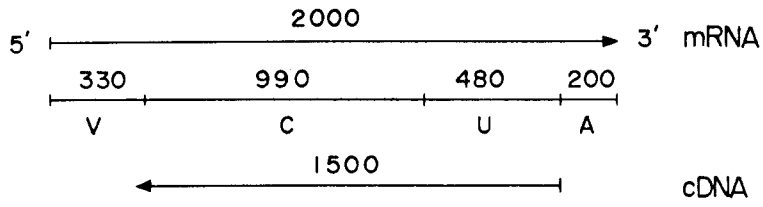


Fig. 1. Diagramatic representation of H chain mRNA and its cDNA. V, variable region sequence; C, constant region sequence; U, untranslated sequence; A, poly(A). Numbers in bases.

III REITERATION FREQUENCY OF THE γ1 GENE

MOPC 31C ^{3}H-cDNA was hybridized to a vast excess of cellular DNA derived from MOPC 31C myeloma. Reassociation of mouse unique DNA was assessed by the addition of a trace amount of ^{14}C-labeled MOPC 31C DNA. All the Cot analyses were carried out in similar double-labeled experiments under the

identical conditions.

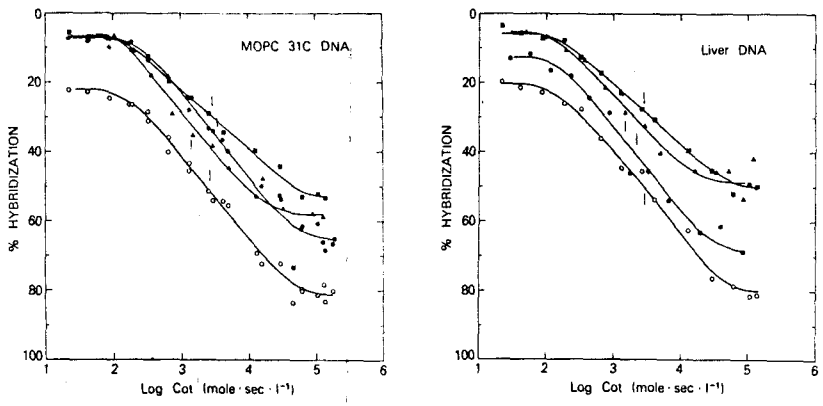


Fig. 2. Hybridization kinetic analyses of MOPC 31C cDNA with total cellular DNA from MOPC 31C myeloma and mouse liver. Arrows indicate Cot1/2 points. Reactions were carried out as described previously (7). The concentrations of MOPC 31C ^{3}H)-cDNA, cellular ^{14}C-DNA and Na^{+} were 465 pg/ml, 9.4 mg/ml and 1 M, respectively. Specific radioactivities of MOPC 31C (^{3}H)-cDNA and ^{14}C-DNA were 1.6 x 10^{7} cpm/μg and 610 cpm/mg, respectively. ● , *(^{3}H)-cDNA;* ○ , *^{14}C-DNA;* ■ , *Kappa chain (^{3}H)-cDNA;* ▲ , *globin (^{3}H)-cDNA.*

As shown in Fig. 2, γ1 cDNA hybridized to about 65% with a Cot1/2 value of 3,500, while that of mouse unique sequences is 2,700. Comparison of these Cot1/2 values yields a reiteration frequency of 0.8 copy per haploid genome for the γ1 gene sequence in MOPC 31C DNA. As control experiments kappa L chain cDNA derived from MOPC 41 mRNA (2) and globin cDNA were hybridized to the same batch of MOPC 31C DNA. Cot1/2 values for the kappa chain cDNA and globin cDNA were 2,900 and 1,400, respectively, indicating that the kappa chain and globin genes are represented 0.93 and 1.9 times per haploid genome respectively, which are in agreement with previous reports (1-9). Similar experiments were carried out with liver DNA. The results are essentially identical to what obtained with MOPC 31C DNA. Namely, reiteration frequencies of the γ1, κ chain and globin genes are calculated to be 1.3,

1.0 and 1.9 copies per haploid genome, respectively. As summarized in Table I, the γ1 gene sequence is present in one copy per haploid genome in IgG1-, IgM- and IgG3-producing myelomas as well as normal mouse tissues.

Table I Reiteration Frequency of the γ1 Gene

Origin of DNA		*Copies/haploid*		
Myeloma	*Subclass*	*Hγ1*	*Lκ*	*Globin*
MOPC 31C	γ_1	*0.8*	*0.93*	*1.9*
MC 101	γ_1	*1.2*		
MOPC 70A	γ_1	*0.8*		
MOPC 104E	μ	*1.0*		
J606	γ_3	*1.1*		
Liver	-	*1.3*	*1.0*	*1.9*
Spleen	-	*1.0*		
Newborn Mouse	-	*0.96*	*1.0*	
Kidney	-	*0.96*		

On the other hand when the γ1 cDNA was hybridized to DNA derived from a IgG2b-producing myeloma, MPC 11, hybridization kinetics are quite different from those shown in Fig. 2.

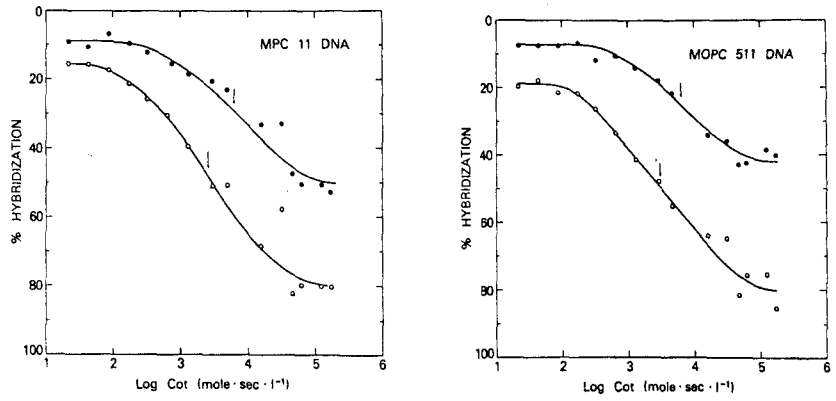


Fig. 3. Hybridization kinetic analyses of MOPC 31C cDNA with total cellular DNA from MPC 11 and MOPC 511 myelomas. ●, ^{3}*H-cDNA;* ○, ^{14}*C-DNA.*

As shown in Fig. 3, a Cotl/2 value of 6,200 was obtained as compared to that of 2,600 for mouse unique copy DNA. The results yield a reiteration frequency of 0.42 copy per haploid genome. The extent of hybridization reached approximately 48%, which is smaller than the value (65%) obtained with MOPC 31C and liver DNA.

We have tested the congruity of the cDNA-DNA hybrids by thermal stability to S1 nuclease digestion. The γ1 cDNA hybrids formed with MOPC 31C DNA and MPC 11 DNA showed sharp melting profiles with Tm values of 90.6° and 86.8°, respectively, indicating that congruently matched duplexes were formed. A slight reduction in a Tm value for MPC 11 DNA hybrids is probably due to degradation of cDNA before hybrids are formed.

The possibility that the reduction of the extent of hybridization is due to a partial deletion of the γ1 sequence was excluded since the experiments using the cDNA preparations of different length (500 bases, 1,000 bases and 1,500 bases) gave the identical extent of hybridization. The γ1 gene sequence seems to be present also in 0.5 copy per haploid genome in a IgA-producing tumor, MOPC 511.

In seven lines of myeloma producing IgG2a, IgG2b, IgA or no H chain the number of γ1 gene copy ranged from 0.39 to 0.56 as summarized in Table II.

Table II Reduction of the Number of the γ1 Gene Copy in Myelomas Producing IgG2a, IgG2b and IgA

Origin of DNA		*Copies/haploid*		
Myeloma	*Subclass*	*H*	*L*	*Globin*
RPC 5	*γ2a*	*0.56*		
MPC 11	*γ2b*	*0.42*	*1.1*	
NP 2	*γ2b*	*0.39*	*1.1*	
MOPC 511	α	*0.48*		
MOPC 315	α	*0.52*		
MOPC 41	-	*0.53*	*1.0*	*1.8*
RPC 20	-	*0.44*		

The maximal extent of hybridization is approximately 45% in average. The results indicate that the average number of the γ1 gene copy is reduced to 0.5 per haploid genome in these myelomas.

Obviously the results can be interpreted in two ways. One possibility is that general chromosomal abnormalities in these myelomas happen to have reduced the apparent reiteration frequency of the γ1 gene (17). The alternative possibility, which we prefer, is that the γ1 gene is specifically deleted from one of the alleles in mouse myelomas which produce the immunoglobulins of defined classes, i.e., IgG2a, IgG2b and IgA.

IV REITERATION FREQUENCY OF THE γ3 GENE

In order to test our favorite possibility mentioned above we have assessed the number of times the gene sequence corresponding the γ3 cDNA probe is represented in various myelomas by hybridization kinetic analysis.

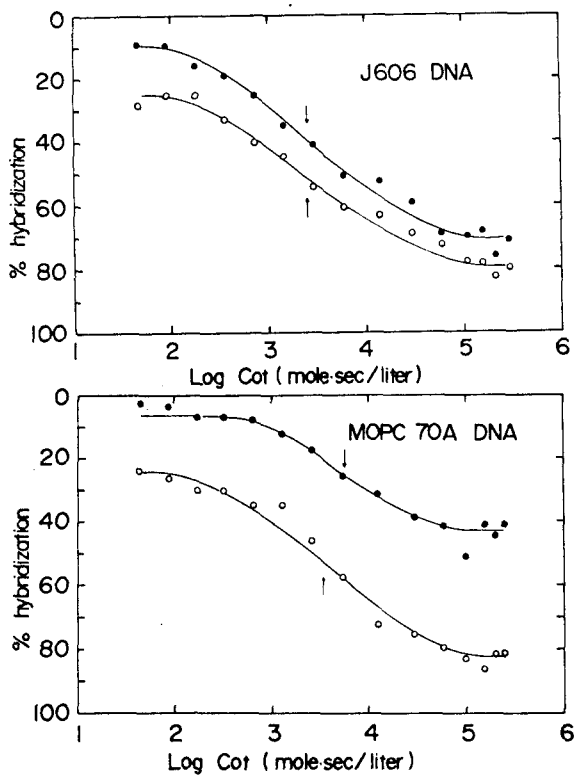


Fig. 4. Hybridization kinetic analyses of J606 cDNA with total cellular DNA from J606 and MOPC 70A myelomas.
● , *^{3}H-cDNA;* ○ , *^{14}C-DNA.*

Typical examples of such experiments are shown in Fig. 4, in which J606 ^{3}H-cDNA was hybridized to a vast excess of cellular DNA derived from J606 (IgG3) and MOPC 70A (IgG1) myelomas, respectively. The results are quite clear. The Cot1/2 values for J606 DNA and MOPC 70A DNA are 3,000 and 5,800 respectively, as compared to a Cot1/2 value of 3,000 for unique copy mouse DNA. These values are equivalent to reiteration frequencies of 1.0 and 0:52 per haploid genome for J606 and MOPC 70A DNA, respectively.

Similar experiments are carried out with DNA derived from seven other myelomas and three tissues, the results of which are summarized in Table III.

Table III Reiteration Frequency of the γ3 Gene

Origin of DNA		*Copy/haploid*
Myeloma	*Subclass*	
MOPC 104E	μ	*0.85*
J606	γ3	*1.0*
MOPC 31C	γ1	*0.55*
MOPC 70A	γ1	*0.52*
MPC 11	γ2b	*0.52*
NP 2	γ2b	*0.5*
UPC 10	γ2a	*0.52*
MOPC 511	α	*0.56*
MOPC 315	α	*0.4*
Liver	-	*1.18*
Kidney	-	*0.88*
Newborn mice	-	*1.27*

IgM- and IgG3-producing myelomas in addition to normal tissues seem to contain one copy of the γ3 gene sequence per haploid genome, whereas IgG1-, IgG2a-, IgG2b- and IgA- producing myelomas have only a half copy of the γ3 gene sequence per haploid genome. It is worth emphasizing that IgG1-producing myelomas, MOPC 31C and MOPC 70A have lost a half copy of the γ3 gene sequence per haploid genome. Nevertheless they retain one copy of the γ1 gene sequence per haploid. Since both sequences are shown to cluster on a small segment of one

chromosome, it is unlikely that the reduction of the number of the γ3 gene copy is attributable to general chromosomal abnormalities in these myelomas. Therefore, we are inclined to conclude that reduction of the γ1 or γ3 gene copy is due to the specific deletion of the particular genes.

Given the reiteration frequency of the γ1 and γ3 genes, we can classify mouse myelomas into three groups. The group I contains one copy each of the γ1 and γ3 gene sequences. The IgM- and IgG3-producers belong to this group. The group II contains half a copy of the γ3 sequence and one copy of γ1 sequence as is the case of the IgG1 producer. The group III consists of IgG2a, IgG2b and IgA producers in which both γ1 and γ3 sequences are represented only 0.5 time per haploid genome.

V THE ALLELIC DELETION MODEL

In order to put all these results together, we propose a model which we call the allelic deletion model. We assume that the specific deletion of C_H gene sequences accompanies the V-C gene recombination which takes place only on one of the alleles as shown in Fig. 5.

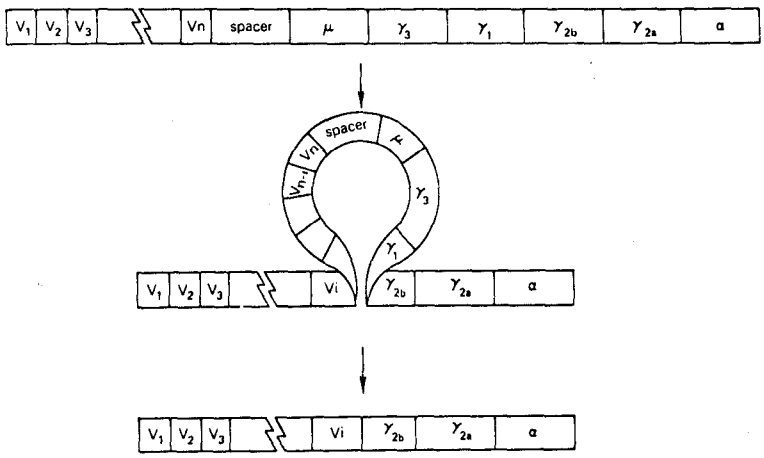


Fig. 5. Allelic deletion model for V_H-V_C recombination. The diagram shows recombination of a V_i sequence with the γ2b gene, resulting in deletion of the μ, γ3 and γ1 genes.

Upon joining of the V_H and C_H gene sequences the portion of the chromosome which resides between two sequences is excised out of the chromosome. Consequently the deletion of the γ1 or γ3 gene in mouse myeloma of the specific class reflects the relative locations of the class-specific C_H genes that have undergone recombination with a V_H gene. Inasmuch as the group I myelomas (IgM and IgG3 producers) retain both γ1 and γ3 gene sequences, the μ and γ3 genes must reside closer to V_H genes than the γ1 gene. It is also reasonable to put the μ gene closer to the V_H sequences than the γ3 sequence. Otherwise, the γ3 gene should be deleted in the IgM-producing myeloma. Obviously the group II gene, γ1 is the second closest to the V_H genes and the group III genes are the furthest. Our preliminary data suggest that the γ2b gene is closer to V_H genes than the γ2a gene. We put the α gene tentatively at the furthest end based on the evolutional consideration. This order of C_H genes is compatible with recent genetic studies using rabbit allotypes (18-20).

We have recently succeeded in purification of γ2a, γ2b and μ chain mRNA from mouse myeloma. Using these probes we will be able to assess whether other C_H gene sequences are deleted according to the manner which the allelic deletion model predicts.

Although no experimental data were available, all the possible mechanisms have been so far proposed to account for the V-C recombination, which include 1) the copy-insertion model (21); 2) the translocation model (22); 3) the inversion model (23) and 4) the looping-out excision model (24). None of the models except the looping-out excision model can explain our results. As a matter of fact we learned this theory very recently and were impressed by its great insight. Kabat's theory was designed to explain mutual exclusivity of gene function in the hemoglobin system and suggested to be applicable to immunoglobulin genes as well.

VI CONCORDANCE OF THE ALLELIC DELETION MODEL WITH IMMUNOLOGICAL PHENOMENA

Any model for the immunoglobulin genes has to explain such unique immunological observations as allelic exclusion and switch of the immunoglobulin class synthesized in a single lymphocyte.

A. Allelic Exclusion

Genetic studies using allotype markers have shown that only one of the alleles of immunoglobulin genes is expressed

in activated B lymphocytes (25-27). Genetic information for the V_H and C_H which are genetically linked on one parental chromosome are coordinately expressed, which is referred to as cis expression (28). Apparently the allelic deletion model postulates allelic exclusion and cis expression by nature.

B. Switch of Immunoglobulin Class

The switch within a clone from μ chain synthesis to γ chain synthesis is well established as the normal process in which B lymphocytes begin to secret IgG and this easily explains the simultaneous presence of both IgM and IgG on the cell surface (29-31). It is not clear, however, whether this switch may include a period in which a single cell simultaneously synthesizes μ and γ chain and, in particular, μ and γ mRNAs. According to the allelic deletion model the successive recombination of a V_H gene sequence first with μ gene then with γ gene can account for the shift of the immunoglobulin class synthesized in a lymphocyte. The switch to the reverse direction has not been found and is very difficult to explain by this model. In view of the relatively long half life of immunoglobulin proteins and mRNAs (32,33), we do not have to postulate the concurrent transcription of μ and γ gene sequences within a single lymphocyte.

VII ACKNOWLEDGEMENTS

We are grateful to Drs. M. Ono and M. Kawakami for purification of H chain mRNA; to Dr. Y. Mano for his encouragement; to Dr. M. Potter for supplying many lines of mouse myeloma and to Drs. P. Leder, J. Gruber and J. W. Beard for a generous gift of reverse transcriptase. This investigation was supported by grants from Jane Coffin Child Memorial Fund for Cancer Reseach, Matsunaga Science Foundation and Ministry of Education, Science and Culture of Japan.

VIII REFERENCES

1. Faust, C. H., Diggelman, H. and Mach, B., Proc. Natl. Acad. Sci., 71, 2491 (1974).
2. Honjo, T., Packman, S., Swan, D. and Leder, P., Proc. Natl. Acad. Sci., 71, 3659 (1974).
3. Stavnezer, J., Huang, R. C. C., Stavnezer, E. and Bishop,J.M.,J. Mol.Biol., 88, 43 (1974).
4. Tonegawa, S., Steinberg, C., Dube, S. and Bernardini, A., Proc. Natl. Acad. Sci., 71, 4027 (1974).

5. Rabbits, T. H., Jarvis, J. M. and Milstein, C., Cell, 6, 5 (1975).
6. France, M-G., Aellen, M-F., Briand, P-A., Faust, C. H., Vassalli, P. and Mach, B., Proc. Natl. Acad. Sci., 73, 727 (1976).
7. Honjo, T., Packman, S., Swan, D., Nau, M. and Leder, P., Biochemistry, 15, 2780 (1976).
8. Leder, P., Honjo, T., Seidman, J. and Swan, D., Cold Spring Harbor Symp. Quant. Biol., 41, 855 (1976).
9. Tonegawa, S., Proc. Natl. Acad. Sci., 73, 203 (1976).
10. Milstein, C., Brownlee, G. G., Harrison, T. M. and Mathews, M. B., Nature New Biol., 239, 117 (1972).
11. Dreyer, W. J. and Bennett, J. C., Proc. Natl. Acad. Sci., 54, 864 (1965).
12. Premkumar, E., Shoyab, M. and Williamson, A. R., Proc. Natl. Acad. Sci., 71, 99 (1974).
13. Bernardini, A. and Tonegawa, S., FEBS Lett., 41, 73 (1974).
14. Ono, M., Kondo, T, Kawakami, M. and Honjo, T., J. Biochem., 81, 949 (1976).
15. Eichmann, K., Immunogenetics, 2, 491 (1975).

17. Yoshida, T. H., Imai, H. T. and Potter, M., J. Natl. Cancer Inst., 41, 1083 (1968).
18. Mage, R. G., Young-Cooper, G. O. and Alexander, C., Nature, New Biol., 230, 63 (1971).
19. Kindt, T. J. and Mandy, W. J., J. Immunol., 108, 1110 (1972).
20. Hamers-Casterman, C. and Hamers, R., Immunogenetics, 2, 597 (1975).
21. Dreyer, W. J., Gray, W. R. and Hood, L., Cold Spring Harbor Symp. Quant. Biol., 32, 353 (1968).
22. Gally, J. A. and Edelman, G. M., Nature, 227, 341 (1970)
23. Tonegawa, S., Hozumi, N., Matthyssens, G. and Schuller, R., Cold Spring Harbor Symp. Quant. Biol., 41, 877 (1976).
24. Kabat, D., Science, 175, 134 (1972).
25. Pernis, B. G., Chiappino, G., Kelus, A. S. and Gell, B. G. H., J. Exp. Med., 122, 853 (1965).
26. Cebra, J. J., Colberg, J. E., Dray, S., J. Exp. Med. 123, 547 (1966).
27. Mäkelä, O. and Cross, A. W., Prog. Allergy, 14, 145 (1970).
28. Knight, K. L. and Hanly, W. C., in "Contemporary Topics in Molecular Immunology" (F.P.Inman and W.J.Mandy, Eds.), Vol. 4., p. 55, Plenum Press, New York, 1975.
29. Kincade, P. W., Lawton, A. R., Bockman, D. E. and Cooper, M. D., Proc. Natl. Acad. Sci., 67, 1918 (1970).

30. Pierce, C. W., Solliday, S. M. and Asofsky, R., J. Exp. Med., 135, 675 (1972).
31. Jones, P. P., Tacier-Eugster, H. and Herzenberg, L. A., Ann. Immunol. (Inst. Pasteur), 125C, 271 (1974).
32. Cowan, N. J. and Milstein, C., J. Mol. Biol., 82, 469 (1974).
33. Melchers, F., Lafleur, L. and Andersson, J. in "Control of Proliferation in Animal Cells" (Clarkson, B. and Baserga, R., Eds.), p. 393, Cold Spring Harbor Laboratory, New York, 1974.

[Reprinted from Biochemistry, (1979) 18, 490.]

Mutual Homology of Mouse Immunoglobulin γ-Chain Gene Sequences†

Yuriko Yamawaki-Kataoka, Kazue Sato, Akira Shimizu, Tohru Kataoka, Yoshitake Mano, Masao Ono, Masaya Kawakami, and Tasuku Honjo*

ABSTRACT: We have assessed the relative homology of mouse immunoglobulin heavy-chain gene sequences using complementary DNAs (cDNAs) synthesized against γ-chain mRNAs (γ1, γ2a, γ2b, and γ3) purified from mouse myelomas. cDNAs complementary to the γ-chain mRNAs did not cross-hybridize with the μ- and α-chain mRNAs, whereas they cross-hybridized to significant extents (22–66%) with the γ-chain mRNAs of other subclasses. The heterologous hybrids formed, however, melt at 5–13 °C lower temperatures as compared to the homologous hybrids, indicating that significant portions of the heterologous hybrids are mismatched. The rates of the cross-hybridization reactions are 2- to 17-fold slower than those of the homologous hybridization reactions. Therefore, the γ-chain gene sequences of four subclasses share a part of homology with each other, but they are different enough to be measured separately. Cross-hybridization analyses indicate that the γ2a and γ2b genes are the most closely related, while the γ1 and γ3 genes are the least related among the γ subclass genes.

Immunoglobulin heavy chains are encoded by a family of V[1] region genes and a set of C region genes (reviewed by Eichmann, 1975). In order to study the organization of immunoglobulin heavy-chain genes, it is essential to isolate the specific probe for each heavy chain gene. Recently we have succeeded in purifying mRNAs encoding mouse immunoglobulin heavy chains (Ono et al., 1977). Using purified mRNAs we have synthesized and purified complementary DNAs (cDNAs) corresponding to the heavy chains of all four γ subclasses, namely, γ1, γ2a, γ2b, and γ3 chains.

In this paper we will report the mutual homology among four γ-chain gene sequences. The extent and rate of hybridization and the thermal stability of hybrids formed showed that although these gene sequences share partial homology with each other, each cDNA is specific to its own subclass.

Materials and Methods

Materials. Mouse myeloma tumors were kindly supplied by Dr. M. Potter of National Institutes of Health, except that MC 101 was provided from Dr. Migita of Kanazawa University. Tumors were maintained as described (Swan et al., 1972). ^{3}H-Labeled dCTP (24 Ci/mmol) was obtained from the Radiochemical Center, Amersham, England, and ^{3}H-labeled dATP and dGTP (12 Ci/mmol) were purchased from New England Nuclear. Hydroxylapatite of DNA grade was obtained from Bio-Rad Lab.

Preparation and Purification of mRNAs and cDNAs. Preparation and purification of heavy-chain mRNAs derived from mouse myelomas MOPC 31C (γ1), HOPC 1 (γ2a), MPC 11 (γ2b), and J 606 (γ3) were done as described in the previous report (Ono et al., 1977). Purities of γ1, γ2a, γ2b, and γ3 mRNAs were 97, 63, 80, and 100%, respectively, as assayed by hybridization kinetic analysis. Detailed characterization of mRNAs will be described elsewhere. [^{3}H]cDNAs complementary to the mRNAs were synthesized using [^{3}H]dCTP, [^{3}H]dATP, and [^{3}H]dGTP by avian myeloblastosis virus reverse transcriptase and purified as described (Honjo et al., 1974). The specific radioactivity of [^{3}H]cDNA was 1.6×10^7 cpm/μg. Although each cDNA showed a single transition in hybridization kinetic analysis to corresponding mRNAs (Ono et al., 1977), γ2a and γ2b cDNAs showed, in our preliminary experiments, the presence of some quantity of contaminants in their preparations. As the contaminants seem to come from minor mRNA species present in common in tumor cells, they were removed by hybridizing γ2a and γ2b

† From the Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan (Y.K., K.S., A.S., T.K., Y.M., and T.H.), and the Molecular Biology Laboratory, School of Medicine, Kitasato University, Sagamihara, Kanagawa 228 (M.O. and M.K.). *Received June 21, 1978.* This study was supported in parts by grants from the Ministry of Education, Science and Culture of Japan, Jane Coffin Child Memorial Fund for Cancer Research, and Matsunaga Science Foundation.

[1] Abbreviations used: V and C regions, variable and constant regions; cDNA, synthetic DNA complementary to mRNA; Cot (Crt) values, product of concentration of nucleotide sequences of DNA (RNA) and time of incubation (mol of nucleotides × s/L).

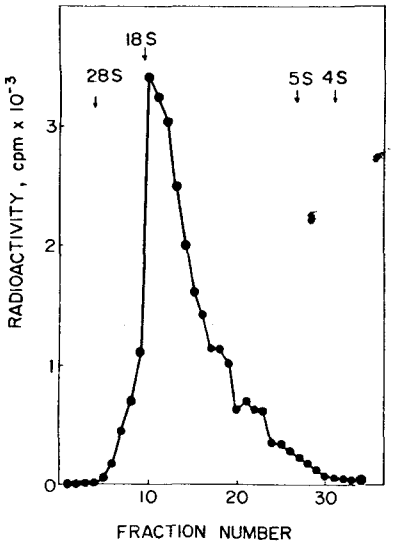


FIGURE 1: Electrophoretic pattern of MPC 11 cDNA. Purified MPC 11 [^{3}H]cDNA was run on polyacrylamide gel electrophoresis in formamide. The reference position of 28S, 18S, 5S, and 4S rRNA from mouse myeloma is also indicated.

cDNA with a crude mRNA preparation derived from a tumor (MOPC 511) producing α chain. After hybridization to a Crt value of 40, the reaction mixture was loaded on a hydroxylapatite column and cDNA remaining as single stranded was eluted with 0.14 M phosphate buffer. By this purification, about 65% of the original cDNA was recovered.

Determination of cDNA Length. [^{3}H]cDNAs were electrophoresed at 130 V for 3 h at room temperature in 4% polyacrylamide gels made up in 99% formamide (Staynov et al., 1972). After electrophoresis the gel was cut into 2-mm thick slices. Each slice was solubilized in 20% H_2O_2 and 20% perchloric acid at 70 °C for 2 h (Mahin & Lofberg, 1966), and then radioactivity was measured by adding 10 mL of Aquasol-2 (New England Nuclear).

cDNA–RNA Hybridization. Hybridization reactions were carried out as described (Honjo et al., 1974). Hybridization was assayed with S1 nuclease digestion as described (Honjo et al., 1976). RNAs were partially purified from various myeloma tumors up to the dT1 stage (Honjo et al., 1974), which comprise about 4% pure mRNA. Briefly, the homogenate of myeloma tumors was centrifuged to remove nuclei and the total cytoplasmic RNA was isolated by phenol—methacresol extraction (Kirby, 1968). mRNA was purified by oligo(dT)-cellulose column chromatography.

Thermal Denaturation of Hybrids. Hybridization was carried out in 0.6 M NaCl–0.2 mM EDTA–20 mM Tris-HCl, pH 7.2, at 75 °C. When a Crt value of 0.5 was attained, the hybridization mixture was immediately chilled and diluted to 0.25 M NaCl. The thermal stability of cDNA–RNA hybrids was determined by measurement of the S1 nuclease resistance upon heating (Honjo et al., 1976).

Results

Characterization of cDNAs. Complementary DNAs were synthesized against γ-chain mRNAs of four subclasses which were highly purified from MOPC 31C (IgG 1), HOPC 1 (IgG 2a), MPC 11 (IgG 2b), and J 606 (IgG 3) myelomas (Ono et al., 1977). The sizes of the [^{3}H]cDNAs were determined by polyacrylamide gel electrophoresis in 99% formamide. A typical electrophoresis pattern is shown in Figure 1. The size distribution patterns were rather heterogeneous with an average chain length of about 1500 nucleotides, having a peak length of 1500, 1900, 1900, and 1600 nucleotides for MOPC

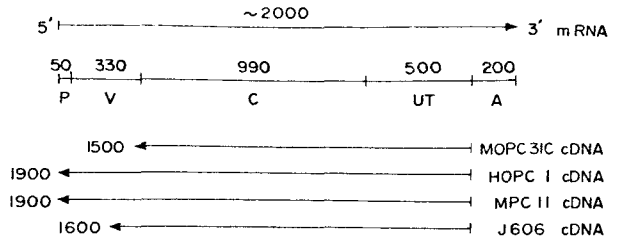


FIGURE 2: Diagramatic representation of γ-class mRNA and cDNAs. P, precursor region sequence determined from the amino acid sequence of the MOPC 315 precursor (Jilka & Pestka, 1977); V, variable region sequence; C, constant region sequence; UT, untranslated sequence; A, poly(A) sequence. Numbers in bases.

31C (γ1), HOPC 1 (γ2a), MPC 11 (γ2b), and J 606 (γ3) cDNA, respectively. γ-Chain mRNAs are 2000 nucleotides long (Ono et al., 1977), of which the V region and C region sequences occupy 330 and 990 nucleotides, respectively. Nucleotide sequence analysis of MOPC 31C cDNA inserted in a plasmid indicates that the untranslated sequence of about 500 nucleotides is present at the 3′ end of mRNA (Honjo and Leder, unpublished observation). Assuming that the mRNAs contain poly(adenylic acid) of 200 nucleotides, these cDNAs include almost the entire C region sequence but do not necessarily include the V region sequence (Figure 2).

Cross-Hybridization within a Subclass. To determine whether these cDNAs extend into the V region sequence or not, the cDNAs were hybridized with partially purified (dT1 stage) mRNAs derived from myelomas producing other γ chains of the same subclass, namely, heavy-chain proteins with the same C region sequence and a different V region sequence.

As shown in Figure 3, MOPC 31C (γ1) cDNA hybridized with MC 101 (γ1) mRNA to the extent similar to that with MOPC 31C mRNA. However, HOPC 1 (γ2a), MPC 11 (γ2b), and J 606 (γ3) cDNAs hybridized to UPC 10 (γ2a), MOPC 141 (γ2b), and FLOPC 21 (γ3) mRNAs, respectively, to less extents than those to the respective homologous mRNAs; the final extents of hybridization obtained for the heterologous mRNAs were about 85 (γ2a) and 90% (γ2b and γ3) of those obtained for the homologous mRNAs. HOPC 1, MPC 11, and J 606 cDNAs seem to be long enough to extend into the V region sequence, while MOPC 31C cDNA covers only the constant region sequence.

The $Crt_{1/2}$ values of these cross-hybridizations are comparable to those of the homologous hybridizations, which indicates the presence of different but roughly similar quantities of mRNA among different tumors. The thermal denaturation profiles of hybrids formed between cDNAs and corresponding partially purified mRNAs were sharp and undistinguishable from those obtained for hybrids with highly purified mRNAs (data not shown).

Cross-Hybridization between γ-Chain Genes of Different Subclasses. The γ-chain cDNAs derived from MOPC 31C (γ1), HOPC 1 (γ2a), MPC 11 (γ2b), and J 606 (γ3) mRNAs were examined for their cross-hybridizability to other mRNA sequences. [^{3}H]cDNA probes were hybridized to large excess of mRNAs from various myelomas and assayed for their extents of cross-hybridization and thermal stability of hybrids formed. All possible sets of cross-hybridizations were examined as shown in Figures 4 and 5. It is clear that, although each cDNA probe reacted extensively (89–94%) with mRNA of the identical subclass, it hybridized poorly with mRNAs of other subclasses and classes under the conditions used. Between the γ-chain subclasses, only small extents (20–39%) of cross-hybridization were observed except in the case of the γ2a–γ2b cross-hybridization where the extent of cross-hy-

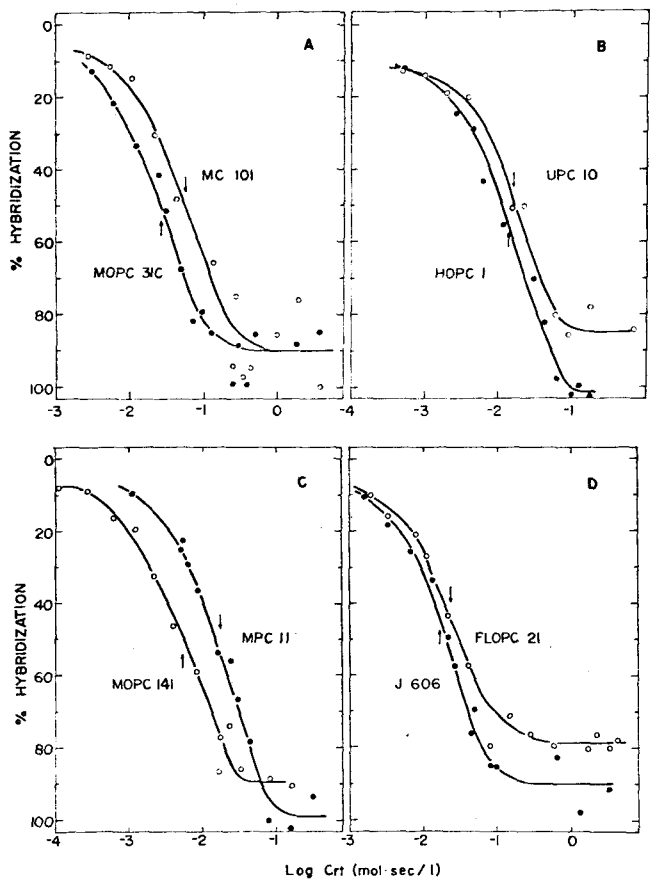


FIGURE 3: Cross-hybridization of γ1, γ2a, γ2b, and γ3 cDNAs with mRNAs of the corresponding subclasses. Hybridization reactions were carried out as described under Materials and Methods. A, MOPC 31C cDNA (γ1) with γ1 mRNAs from MOPC 31C and MC 101 myelomas. B, HOPC 1 cDNA (γ2a) with γ2a mRNAs from HOPC 1 and UPC 10 myelomas. C, MPC 11 cDNA (γ2b) with γ2b mRNAs from MPC 11 and MOPC 141 myelomas. D, J 606 cDNA (γ3) with γ3 mRNAs from J 606 and FLOPC 21 myelomas. Arrows indicate $Crt_{1/2}$ points. $Crt_{1/2}$ values for mRNAs are 2.7×10^{-2} (MOPC 31C), 5.6×10^{-2} (MC 101), 1.4×10^{-2} (HOPC 1), 1.6×10^{-2} (UPC 10), 1.8×10^{-2} (MPC 11), 5.5×10^{-3} (MOPC 141), 1.65×10^{-2} (J 606), and 2.4×10^{-2} (FLOPC 21).

bridization reached as high as 60% of the input cDNA. The possibility that the small extents of cross-hybridization are due to contaminants present in the cDNA probes was excluded by the experiments in which the γ1 gene cloned in a phage did hybridize to similar extents with all the other γ-chain mRNAs (Honjo, unpublished data). α- and μ-chain mRNAs did not hybridize to the γ-chain probes at any significant degree even at a Crt value of 10 or 100. Neither did liver mRNA hybridize with any of the γ-class probes. The results indicate that the γ-chain cDNAs used are not contaminated with other sequences present in myelomas or liver.

Since a $Crt_{1/2}$ value is inversely proportional to a rate of hybridization reaction, relative rates of cross-hybridization of a single mRNA to different cDNAs were calculated from $Crt_{1/2}$ values obtained in Figures 4 and 5. Table I shows that the heterologous hybridization reactions are 2- to 17-fold slower than the homologous reactions. It seems that the smaller the extent of cross-hybridization, the slower the reaction rate. For example, J 606 mRNA hybridized with MOPC 31C cDNA to 26% with a 17-fold slower reaction rate as compared to that with J 606 cDNA. On the other hand, HOPC 1 mRNA hybridized with MPC 11 cDNA to 60%, with a twofold slower rate than that with HOPC 1 cDNA. Retardation in a hybridization rate has been reported for imperfectly matched DNA–DNA or mRNA–cDNA hy-

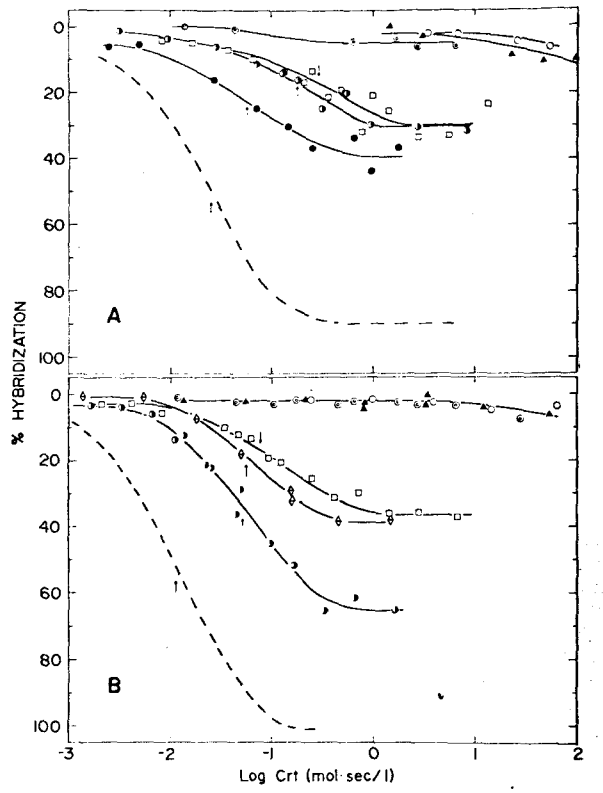


FIGURE 4: Cross-hybridization of γ1 and γ2a cDNAs with mRNAs of other subclasses. Hybridization reactions were carried out as described under Materials and Methods. Arrows indicate $Crt_{1/2}$ points. $Crt_{1/2}$ values are shown in Table I. Hybridization of each cDNA with homologous mRNA is also shown on these figures in dotted line. A, MOPC 31C cDNA (γ1). B, HOPC 1 cDNA (γ2a). ◈, MOPC 31C mRNA (γ1); ●, HOPC 1 mRNA (γ2a); ◑, MPC 11 mRNA (γ2b); □, J 606 mRNA (γ3); ▲, MOPC 104E mRNA (μ); ⊙, MOPC 511 mRNA (α); ○, liver mRNA.

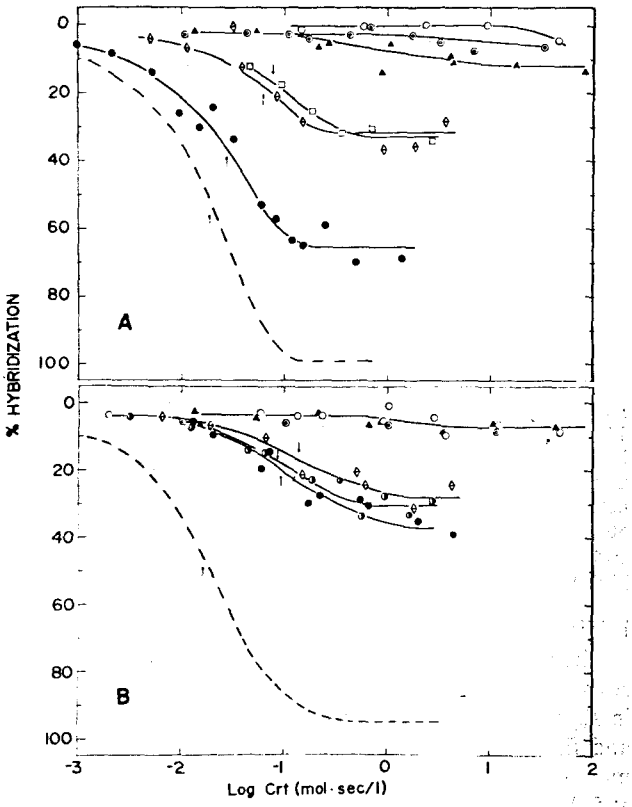


FIGURE 5: Cross-hybridization of γ2b and γ3 cDNAs with mRNAs of other subclasses. Condition of hybridization and explanations of symbols are described in the legend of Figure 4. A, MPC 11 cDNA (γ2b); B, J 606 cDNA (γ3).

Table I: Summary of Cross-Hybridization

mRNA		cDNA MOPC 31C	HOPC 1	MPC 11	J 606
MOPC 31C	$Crt_{1/2} \times 10^2$	2.7	5.9	6.1	14
	rate[a]	1	0.46	0.44	0.19
	extent[b] (%)	100	41	29	22
	ΔT_m[c]		10.4	10.1	12.9
HOPC 1	$Crt_{1/2} \times 10^2$	8.9	1.4	2.8	9.1
	rate[a]	0.16	1	0.5	0.15
	extent[b] (%)	43	100	63	37
	ΔT_m[c]	11.5		5.2	6.4
MPC 11	$Crt_{1/2} \times 10^2$	19	5.4	1.8	13
	rate[a]	0.10	0.33	1	0.14
	extent[b] (%)	32	66	100	30
	ΔT_m[c]	11.6	5.7		8.9
J 606	$Crt_{1/2} \times 10^2$	30	7.9	8.4	1.65
	rate[a]	0.06	0.21	0.20	1
	extent[b] (%)	29	37	31	100
	ΔT_m[c]	13.0	6.4	9.7	

[a] The relative rate of hybridization of a single mRNA to different cDNAs is shown by taking the rate of the homologous hybridization as 1. [b] The relative extent of hybridization of a single mRNA to different cDNAs is normalized by taking that of the homologous reaction as 100%; the actual final extents for the homologous reactions varied from 89 to 94%. [c] Difference of T_m values of the heterologous hybrids formed between a single cDNA and different mRNAs from that of the homologous hybrids. T_m values for the homologous hybrids of MOPC 31C, HOPC 1, MPC 11, and J 606 cDNAs are 93, 90, 92, and 94 °C, respectively.

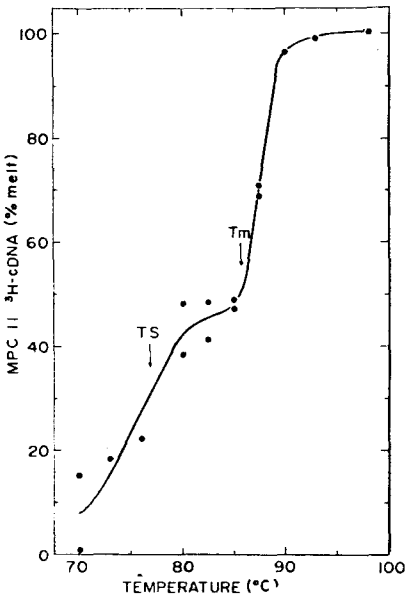


FIGURE 6: Thermal stability profile of hybrid formed between MPC 11 cDNA (γ2b) and HOPC 1 mRNA (γ2a). MPC 11 cDNA was hybridized with HOPC 1 mRNA to a Crt value of 0.5 as described under Materials and Methods. The final extent of hybridization was 65%. The temperature at which a half of the hybrids melt (T_m) was 86 °C, and the temperature at which a half of total hybridizable cDNA (see text) remained as hybrid form (TS) was 78 °C. Arrows indicate positions of T_m and TS.

bridization (Kohne et al., 1972; Leder et al., 1973).

The thermal stability of hybrids formed was measured by resistance to S1 nuclease digestion (Figure 6). Differences in T_m values between the homologous and heterologous hybrids and final extents of the cross-hybridization are also summarized in Table I. The heterologous hybrids formed melt at 5–13 °C lower temperatures as compared to the homologous hybrids, indicating that significant portions of the heterologous hybrids are not congruently base-paired. A ΔT_m value and extent of hybridization for a given combination of mRNA and cDNA were, as expected, in good agreement with those for the reciprocal combination of cDNA and mRNA.

Table II: Relative Homologies between γ-Chain Gene Sequences

gene sequences	ΔTS[a]	% divergence[b]
γ1–γ2a	25	38
γ1–γ2b	30	45
γ1–γ3	42	63
γ2a–γ2b	13	20
γ2a–γ3	25	38
γ2b–γ3	30	45

[a] Difference of thermal stability for total hybridizable cDNA used between the cross-hybrid and the homologous hybrid. [b] Calculated by ΔTS × (1.5% nucleotide pairs substituted/1 °C lowering of TS) according to Kohne et al. (1972).

From the extent of hybridization and the thermal stability of hybrids, it is possible to estimate the evolutional proximity among the γ genes. The temperature (T_m) at which a half of the hybrids melt is used often as a measure of divergence in nucleotide sequences. However, a T_m value of hybrid depends on the temperature at which hybrids are formed. Thus we used the temperature (TS) at which 50% of the total hybridizable cDNA remains in a hybrid form (Kohne et al., 1972) to estimate the divergence of nucleotide sequences among γ-chain genes. The total hybridizable cDNA represents the fraction of C region sequence which varied from 85 (γ2a) to 100% (γ1) of cDNAs (Figure 3). Table II summarizes ΔTS values of cross-hybrids and percentages of nucleotide substitution calculated. It is evident from this table that γ2a and γ2b gene sequences are most closely related, showing 80% homology in nucleotide sequences. On the other hand, 63% of nucleotide sequences seems to be substituted between γ1 and γ3 chain genes, indicating that the γ1 and γ3 genes are most distantly related among the γ gene subclasses. A close relationship between γ2a and γ2b chain is also reported based on serological studies and tryptic peptide maps of the Fc fragment of these proteins (Potter, 1972).

Discussion

We have shown that the γ-chain genes are different from each other although they share partial homology in nucleotide sequence. The divergence in nucleotide sequence is manifested as the low extent of cross-hybridization, reduced thermal stability of heterologous hybrids, and retardation in a rate of cross-hybridization (Figures 4 and 5 and Table I). These properties enable us to distinguish the γ-chain gene sequences from one another distinctively. Class specificity of the γ-chain cDNAs is certified by the fact that none of them cross-hybridized with the μ- and α-chain mRNA.

The specificity of the γ-chain cDNAs offers a great virtue to quantitate the γ gene sequences represented in mouse genome. A relatively small extent and a decreased rate of cross-hybridization indicate that, in the case of Cot analysis, the formation of homologous hybrids is predominant and completes almost before the formation of heterologous hybrids. Each γ-chain cDNA has little chance to form heterologous hybrids since reassociation of the γ-chain gene DNA proceeds faster, leaving little heterologous γ gene sequences available for hybridization with the cDNA. Taking all these into consideration, we assume that the contribution of cross-hybridization to the Cot curve would be negligible. This assumption has been verified by sharp melting profiles of cDNA–DNA hybrids and monophasic unique kinetic curves (Honjo & Kataoka, 1978). Although the γ1, γ2a, γ2b, and γ3 genes are shown to present one copy each per haploid genome in mouse normal tissues, the allelic deletion of these

genes was reported in mouse myelomas (Honjo & Kataoka, 1978).

Since the γ-chain gene sequences did not show any homology with other class (α and μ) gene sequences, early divergence of classes of immunoglobulins must have occurred. This observation agrees with phylogenetic evidence that lower vertebrates such as sharks appear to have only IgM and amphibians have IgM and IgG, although all these classes in addition to IgA commonly exist in mammalians (Fudenberg et al., 1972). The emergence of subclasses is probably a result of recent evolution because the members of the γ-chain gene share considerable homology with each other. Pink et al. (1971) have concluded that subclasses in several species have evolved after differentiation of species. The amino acid sequence studies on γ1 chain (Milstein et al., 1974) and γ2a chain (Fougereau et al., 1976), which are the only two cases that the complete amino acid sequences of mouse immunoglobulin heavy chains are determined, showed that 62% of the C region sequences of two chains is homologous. Our estimation that nucleotide sequence homology between the γ1 and γ2a genes is about 62% (Table II) is in agreement with the amino acid sequence data.

Acknowledgments

We thank Dr. M. Potter for his kind supply of the myeloma tumors and Dr. P. Leder for a generous gift of reverse transcriptase.

References

Eichmann, K. (1975) *Immunogenetics (N.Y.) 2*, 491.

Fougereau, M., Bourgois, A., de Preval, C., Rocca-Serra, J., & Schiff, C. (1976) *Ann. Immunol. (Paris) 127*, 607.

Fudenberg, H. H., Pink, J. R. L., Stites, D. P., & Wang, A. C. (1972) in *Basic Immunogenetics*, p 15, Oxford University Press, New York.

Honjo, T., & Kataoka, T. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2140.

Honjo, T., Packman, S., Swan, D., Nau, M., & Leder, P. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 3659.

Honjo, T., Packman, S., Swan, D., & Leder, P. (1976) *Biochemistry 15*, 2780.

Jilka, R. L., & Pestka, S. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5692.

Kirby, K. S. (1968) *Methods Enzymol. 12B*, 87.

Kohne, D., Chiscon, J., & Hoyer, B. (1972) *J. Hum. Evol. 1*, 627.

Leder, P., Ross, J., Gielen, J., Packman, S., Ikawa, Y., Aviv, H., & Swan, D. (1973) *Cold Spring Harbor Symp. Quant. Biol. 38*, 753.

Mahin, D. T., & Lofberg, R. T. (1966) *Anal. Biochem. 16*, 500.

Milstein, C., Adetugbo, K., Cowan, N., & Secher, D. (1974) Prog. *Immunol., Proc. Int. Congr. Immunol., 2nd*, 157.

Ono, M., Kondo, T., Kawakami, M., & Honjo, T. (1977) *J. Biochem. (Tokyo) 81*, 949.

Pink, R., Wang, A. C., & Fudenberg, H. H. (1971) *Annu. Rev. Med. 22*, 145.

Potter, M. (1972) *Physiol. Rev. 5*, 631.

Staynov, D. Z., Pinder, J. C., & Gratzer, W. B. (1972) *Nature (London), New Biol. 235*, 108.

Swan, D., Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1967.

126. *Purification and Cloning of a Ribosomal RNA Gene Fragment from Mouse DNA*

By Yukio MISHIMA,*) Masaharu SAKAI,*) Masami MURAMATSU,*) Tohru KATAOKA,**) and Tasuku HONJO**)

(Communicated by Toshio KUROKAWA, M. J. A., Dec. 12, 1978)

Introduction. Since the advent of the cloning systems employing either plasmids or λ phage, the isolation and characterization of a single gene from a eukaryotic genome has become a potent approach to the understanding of the complex mechanism of gene regulation in mammalian cells. Several eukaryotic genes from various species have already been isolated and characterized at least partially. We have also reported cloning of an immunoglobulin heavy chain gene from the newborn mouse (Honjo, 1978; Honjo, Kataoka, and Y-Kataoka, in preparation).

The genes coding for ribosomal RNA (rRNA) have been cloned from Yeast (Kramer *et al.*, 1976), Drosophila (Thomas *et al.*, 1974), Xenopus (Morrow *et al.*, 1974), Mouse (Tiemeier *et al.*, 1977) and Chicken (McClements and Skalka, 1977). For mammalian rRNA genes, however, the fragments so far obtained were only internal segments containing neither the promoter nor the terminator region. In order to study the regulatory mechanism of rRNA transcription, it is essential to isolate a fragment containing the 5′-terminal region of the rRNA gene which would contain the promoter.

This work has been initiated for this purpose and, in this report, we describe a successful isolation of such a gene fragment from mouse DNA.

Materials and methods. High molecular weight DNA from newborn mice was prepared by a phenol-chloroform method as described elsewhere (Honjo *et al.*, in preparation). The digestion of the mouse DNA with EcoRI, RPC-5 column chromatography and agarose gel electrophoresis were also described in the paper cited above. The detection of 18S rRNA sequences was performed as described (Honjo *et al.*, in preparation). The procedures of ligation and phage DNA preparation were as described by Tiemier *et al.* (1977). Introduction of DNA by *in vitro* packaging technique was carried out as

*) Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo.

**) Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo.

described (Becker and Gold, 1975; Hohn and Murray, 1977; Sternberg *et al.*, 1977; Blattner, personal communication). Detection of the recombinant phage by *in situ* hybridization was essentially according to Benton and Davis (1977). Detection of the recombinant plasmid by *in situ* colony hybridization was done as described by Grunstein and Hogness (1975). Transformation of Escherichia coli LE392, a derivative of strain ED8656, by pBR322 ligated with rRNA gene fragments was done according to Enea *et al.* (1975). ^{32}P- or ^{3}H-labeled complementary DNA (cDNA) of 18S rRNA was prepared by reverse transcription with random oligonucleotide as described by Taylor *et al.* (1976). ^{3}H-cDNA has been purified by incubating to a Cot of 10^{-2} and fractionating by a hydroxyapatite column. The specific activity of ^{32}P- and ^{3}H-cDNA was $>1.0\times10^{8}$ and 1.0×10^{7} cpm/μg, respectively. All the cloning procedures were carried out in a P-2 laboratory according to the revised NIH Guidelines. This also conformed to the current recommendation by the Japanese Committee on Recombinant DNA.

Results and discussion. 1) *Concentration of ribosomal RNA gene fragments containing the promoter region.* Fig. 1 presents a diagram of the structure of ribosomal RNA genes of the mouse. When the mouse DNA was digested with a restriction endonuclease EcoRI, two categories of gene fragments were produced which could be identified with labeled rRNA probes (Arnheim and Southern, 1977). Although the smaller fragment containing parts of both 18S and 28S RNA was homogeneous in size (~6.6 kb) the larger fragment containing only 18S RNA and the 5′-terminal region of the gene was rather heterogeneous (14–17 kb) suggesting sequence heterogeneity of the spacer region preceding the promoter.

In order to isolate the fragment containing the promoter region, approximately 40 mg of newborn mouse DNA digested with EcoRI was fractionated by an RPC-5 column and each fraction tested for the presence of rRNA sequences after running on an agarose gel.

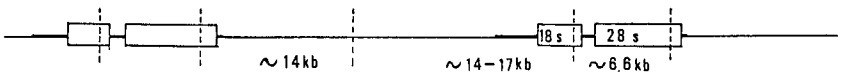


Fig. 1. Organization of ribosomal RNA gene of the mouse. Broken bars indicate the cleavage sites of restriction endonuclease EcoRI. Thick lines indicate the transcribed spacer of 45S-preribosomal RNA. kb: Kilobases.

The fractions containing 14–17 kb bands hydridizable with rRNA probe (data not shown) were pooled and used for cloning. The general scheme for the cloning is summarized in Fig. 2.

Newborn mouse (BALB/c) DNA
↓
EcoRI complete digest
↓
Concentration by RPC-5 column chromatograpy
↓
Cloning with λgt *WES*·λB vector by *in vitro* packaging technique
↓
Plaque detection by filter hybridization
↓
Propagation of positive clones
↓
Transfer to pBR322 plasmid
↓
Clone:pMrEL-1

Fig. 2. Scheme of rDNA Cloning.

2) *Cloning and screening of Ribosomal RNA gene fragments.* The DNA fragments partially purified with respect to ribosomal RNA sequence as described above were ligated *in vitro* with two outer fragments of a mutant coliphage lambda (λgt *WES*·λB) and transfected into a mutant strain of Escherichia coli (LE392). We used the *in vitro* packaging method for the cloning of these DNA fragments,

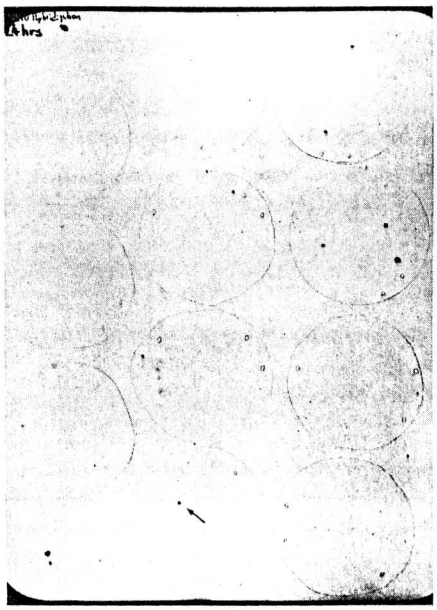


Fig. 3. Detection of hybrid clones. Approximately 2,000 phage clones were transferred to each filter. Detection of 18S rRNA gene was carried out by hybridization with ^{32}P-18S-cDNA followed by autoradiography. The arrow indicates the recombinant clone #3.

Table I. Presence of 18S rRNA sequence in clone #3 DNA

Input (^{3}H-18S-cDNA)	100 %
S_1-Nuclease treated	0
18S rRNA	97.9
#3	47.4
#7	1.7

Percentages of the input counts (564 cpm) were shown. S_1-unclease-treated ^{3}H-cDNA count (45 cpm) was subtracted from each counts.

because it showed a much higher ($>10^2$) transfection efficiency than ordinary Ca^{2+} method in our preliminary experiments.

Approximately 20,000 phage plaques were screened by *in situ* hybridization using ^{32}P-labeled cDNA of 18S rRNA as a probe. Six out of twelve positive spots obtained (Fig. 3) were picked up and separately propagated in a small scale (2 ml) to prepare DNA, which was then tested for the presence of 18S rRNA sequence both by liquid hybridization and by Southern technique after digesting with EcoRI.

Although all the clones were positive with ^{32}P-cDNA of 18S rRNA, one of the clones named #3 was further examined using Cot-purified ^{3}H-cDNA (see Materials and methods) as a probe. Table I shows the clone #3 DNA hybridized with 47.4% of input cDNA, whereas clone #7 DNA which had not hybridized originally, did not hybridize with ^{3}H-cDNA at all. These results indicate that clone #3 DNA does contain nearly half of the 18S rRNA sequence.

3) *Re-cloning of the clone #3 DNA into pBR322.* On examination by gel electrophoresis, clone #3 was found to be contaminated wit hanother clone containing a 7.4 kb fragment of mouse DNA. It has also been known that phage DNA having a large inserted DNA fragment tends to cause deletions during growth. In addition, the ribosomal DNA fragment contained in clone #3 had a size (14–15 kb) which was rather hard to separate from the DEF fragment of λgt *WES* by agarose gel. To solve these problems, we tried to transfer the

Table II. Presence of 18S rRNA sequence in recloned #3 DNA-pBR322-recombinant

Input (^{3}H-18S-cDNA)	100 %
S_1-Nuclease treated	0
18S rRNA	72.3
Clone #(1-4)	48.8

Precentages of the input counts (749 cpm) that were resistant to S_1-nuclease were shown. S_1-nuclease treated ^{3}H-18S-cDNA counts (56 cpm) was subtracted from each counts.

fragment in clone #3 into a plasmid pBR322 DNA in which the inserted DNA was supposed to be more stable and could be recovered more easily.

Clone #3 DNA was digested with EcoRI and the fragments were ligated with pBR322 DNA which had been cleaved with the same enzyme and treated with alkaline phosphatase. Among 60 colonies screened, one positive clone numbered as #(1–4) was obtained. The existence of 18S rRNA sequence was examined by liquid as well as filter hybridization. The results shown in Table II indicate that this clone actually contains roughly half of the 18S rRNA sequence. To demonstrate that the size of the cloned rRNA gene fragment was the same as that originally concentrated and picked up, the recombinant plasmid DNA was digested with EcoRI, electrophoresed on an agarose gel and examined by the Southern blotting technique (1975). As seen in Fig. 4, whereas the DNA in an originally pooled DNA fraction (#30) contained somewhat heterogeneous bands ranging from 14 to 17 kb, the cloned DNA #(1–4) digested with EcoRI showed a single band of approximately 14 kb which was hybridized with cDNA complementary to 18S rRNA. This band did not hybridize with the cDNA

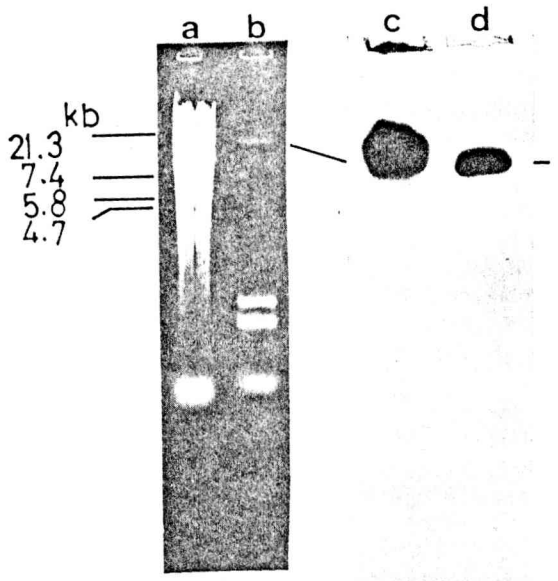


Fig. 4. Presence of 18S rRNA sequence in clone ♯(1–4). (a, b) shows the ethidium bromide-stained 1% agarose gel and (c, d) represents the autoradiography which was hybridized with ^{32}P-18S-cDNA. DNA of columns (a) and (c) was the fraction ♯30 of RPC-5 column chromatography containing the 14–17 kb rDNA fragment. DNA of column (b) and (d) was EcoRI-digested ♯(1–4) plasmid clone. EcoRI fragments of λCI857 were used as size markers.

complementary to 28S rRNA (data not shown). This recombinant plasmid containing a part of 18S rRNA sequence and most probably a promoter region of an rRNA gene of the mouse was designated pMrEL-1.

While this study was under way, we recieved a personal communication that M. H. Edgell at the University of North Carolina had obtained a similar clone from the mouse.

Summary. Newborn mouse DNA was digested with a restriction endonuclease EcoRI and concentrated with respect to ribosomal RNA sequences by an RPC-5 column.

DNA fragments of 14–17 kilobases in length, most probably containing promoter region of the ribosomal RNA gene, were used for cloning with λgt *WES*·λB as a vector using an *in vitro* packaging technique. Several clones containing 18S rRNA sequences were obtained. One of the clones which was transferred to a plasmid pBR322 (designated as pMrEL-1) was 14 kilobases in length, having only a part of 18S rRNA sequence. These results strongly suggest that this fragment carries a promoter region of the ribosomal RNA gene.

Acknowledgements. We thank Dr. P. Leder, NIH, USA, for his generous gift of λgt *WES*·λB-LE392 vector-host system used in this study. We also thank Professor Y. Mano, University of Tokyo, for continuous encouragement. This investigation was supported in part by grants from Ministry of Education, Science and Culture of Japan.

References

Arnheim, N., and Southern, E. M. (1977): Cell, **11**, 363–370.
Becker, A., and Gold, M. (1975): Proc. Natl. Acad. Sci. U.S.A., **72**, 581–585.
Benton, W. D., and Davis, R. W. (1977): Science, **196**, 180–182.
Enea, V., Vovis, G. F., and Zinder, N. D. (1975): J. Mol. Biol., **96**, 495–509.
Grunstein, M., and Hogness, D. S. (1975): Proc. Natl. Acad. Sci. U.S.A., **72**, 3961–3965.
Hohn, B., and Murray, K. (1977): Proc. Natl. Acad. Sci. U.S.A., **74**, 3259–3263.
Honjo, T. (1978): Immunology '78 (eds. Yamamura, Y., and Tada, T.). Nakayama Shoten, Tokyo, pp. 75–82.
Kramer, R., Cameron, J., and Davis, R. (1976): Cell, **8**, 227–232.
McClements, W., and Skalka, A. M. (1977): Science, **196**, 195–197.
Morrow, J. F., Cohen, S. N., Chang, A. C. Y., Boyer, H. W., Goodman, H. M., and Helling, R. B. (1974): Proc. Natl. Acad. Sci. U.S.A., **71**, 1743–1747.
Sternberg, N., Tiemeier, D., and Enquist, L. (1977): Gene, **1**, 255–280.
Southern, E. M. (1975): J. Mol. Biol., **98**, 503–517.
Taylor J. H., Illmensee, R., and Summers, J. (1976): Biochim. Biophys. Acta, **228**, 324–330.
Thomas, M., Cameron, J., and Davis, R. (1974): Proc. Natl. Acad. Sci. U.S.A., **71**, 4579–4583.
Tiemeier, D. C., Tilghman, S. M., and Leder, P. (1977): Gene, **2**, 173–191.

EXISTENCE OF BOTH KAPPA AND LAMBDA LIGHT CHAIN MESSENGER RNA SEQUENCES IN MOUSE MYELOMA, MOPC-104E, KNOWN AS A LAMBDA CHAIN PRODUCER

Masao Ono and Masaya Kawakami

Molecular Biology Laboratory, School of Medicine, Kitasato University, Sagamihara, Japan

Tohru Kataoka and Tasuku Honjo

Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo, Japan

Received December 7,1976

SUMMARY: A mouse myeloma, MOPC-104E, which is known to synthesize and secrete λ type light chain protein as a constituent of immunoglobulin M, was shown to contain mRNA sequences coding for κ as well as λ type light chain protein. Light chain mRNA sequences were quantitated by nucleic acid hybridization reaction using radioactive DNA complementary to light chain mRNAs which had been purified from other myelomas. The amount of κ type light chain mRNA present in MOPC-104E is almost equivalent to that of λ type light chain mRNA. κ chain mRNA was not separated from λ chain mRNA either by centrifugation in sucrose density gradient or by polyacrylamide gel electrophoresis in formamide.

INTRODUCTION: It seems to be generally taken for granted that immunoglobulin mRNA derived from a mouse myeloma encodes the immunoglobulin protein secreted by the myeloma. In fact, this had been proved in most of the myelomas (1-6), although some mRNA had not been well characterized (7,8). MOPC-104E myeloma is known to produce λ type light (L) chain and to secrete it as a constituent of IgM (9). Upon purification of L chain mRNA from MOPC-104E we have found that the myeloma contains κ type L chain mRNA as much as λ type L chain mRNA. The results give warning against using MOPC-104E mRNA or other not-well-characterized myeloma mRNA as probe for analysis of immunoglobulin genes or mRNAs.

MATERIALS AND METHODS

Mouse myeloma cell lines were kindly supplied by Dr. M. Potter, NIH. MOPC-104E (Kyoto) is a cultured cell line donated by Dr. Y. Namba of Kyoto University (10). MOPC-104E (Osaka) was provided by Dr. T. Kishimoto of Osaka University. Purification of L_κ chain mRNAs from MOPC-41 and MOPC-31C myelomas

Abbreviations: cDNA, complementary DNA; Crt value, product of concentration of nucleotide sequences of RNA and time of incubation (mole of nucleotides x sec/liter); $Crt_{1/2}$, Crt value at 50% hybridization; L_κ and L_λ chains, kappa and lambda light chains of immunoglobulin protein

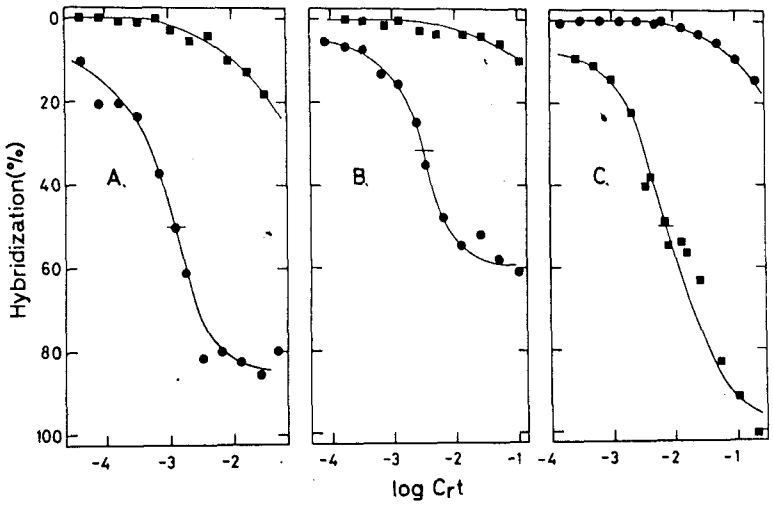


Fig. 1. Cross-hybridization of L_κ and L_λ cDNAs with myeloma RNAs. Total reaction mixture (150µl) contained 12,400cpm (4.1ng) of MOPC-31C L_κ cDNA or 9,650cpm (0.97ng) of RPC-20 L_λ cDNA. Fifteen µl aliquots were removed at appropriate intervals. The RNA was present in at least 15-fold excess over cDNA. A. MOPC-31C immunoprecipitated RNA. B. MOPC-41 RNA. C. RPC-20 RNA. Hybridization with L_κ cDNA (●); L_λ cDNA (■). Endogenous S_1 nuclease resistance (2.6% for L_κ cDNA and 2.0% for L_λ cDNA) was subtracted from all values.

and preparation of L_κ chain cDNA from respective mRNA have been described (1, 11). L_λ chain cDNA was prepared from L_λ chain mRNA purified from RPC-20 myeloma (5,6). Unless otherwise specified, mRNA preparations employed for hybridization were partially purified by two successive chromatographies on an oligo (dT)-cellulose column as described before (1). Total cytoplasmic RNA was obtained according to Kirby's method (13). Hybridization reactions were carried out in 0.6M NaCl at 75°C and assayed by S_1 nuclease digestion as described (1). Other sources of reagents have been described (1,5,6,11,12).

RESULTS

Characterization of hybridization probe

L_κ chain cDNA prepared from MOPC-41 mRNA has been shown to hybridize specifically to L_κ chain mRNA (1,6). Reciprocally, L_λ chain cDNA prepared from RPC-20 mRNA has been demonstrated to be specific for L_λ chain mRNA (1,6). In these myelomas the chain specificity of L chain mRNA present coincides with that of L chain protein secreted by the myeloma. In order to confirm these results we have examined specificity of L_κ chain cDNA prepared from MOPC-31C mRNA (11). As shown in Fig. 1, MOPC-31C cDNA hybridized with L_κ chain mRNAs derived from

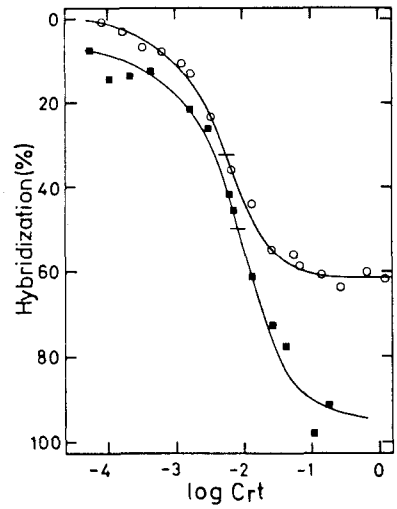


Fig. 2. Hybridization of MOPC-104E (Kyoto) RNA with L_κ and $L\lambda$ cDNA. Hybridization was carried out as described in Fig. 1. Hybridization with L_κ cDNA (O); L_λ cDNA (■).

MOPC-31C and MOPC-41 myelomas. The $Crt_{1/2}$ values (1.2×10^{-3} for MOPC-31C mRNA and 3×10^{-3} for MOPC-41 mRNA) of each hybridization are comparable to that reported for respective mRNA (1,11). The lower extent of hybridization to the heterologous mRNA reflects the proportion of the cDNA complementary to the constant region sequence. In contrast, MOPC-31C cDNA hardly hybridized with L_λ chain mRNA derived from RPC-20 myeloma at Crt values where L_λ chain (RPC-20) cDNA hybridized to completion (Fig. 1C). A slight hybridization observed at high Crt values may reflect a contaminant L_κ chain mRNA sequence present at less than 1/100th concentration of L_λ chain mRNA in RPC-20 myeloma (1). If the slight hybridization is due to a contaminant L_λ cDNA present in MOPC-31C L_κ cDNA, the $Crt_{1/2}$ values for hybridization of RPC-20 mRNA with L_κ (MOPC-31C) and L_λ (RPC-20) cDNAs should be identical.

L_λ chain cDNA prepared from RPC-20 mRNA showed little hybridization with L_κ chain mRNAs at Crt values comparable to hybridization with L_κ chain cDNA (Fig. 1A, B), whereas the L_λ chain cDNA hybridized with homologous mRNA at low $Crt_{1/2}$ value (8×10^{-3})(Fig. 1C). These results clearly indicate that

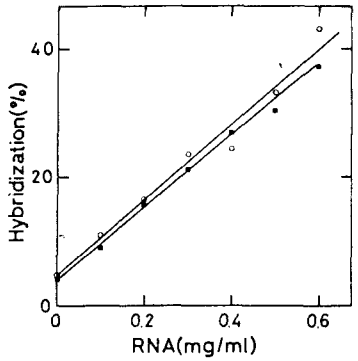


Fig. 3. Hybridization of MOPC-104E (NIH) cytoplasmic RNA with L_κ and L_λ cDNA. Increasing amounts of MOPC-104E cytoplasmic RNA were hybridized with MOPC-41 L_κ cDNA (2,200cpm) and RPC-20 L_λ cDNA (2,400cpm) in 200µl reaction mixture for 36 min. Hybridization with L_κ cDNA (O); L_λ cDNA (■).

MOPC-31C L_κ chain cDNA and RPC-20 L_λ cDNA are specific for the respective chain type.

Presence of both κ and λ chain mRNA sequences in MOPC-104E myeloma

The concentrations of L_κ and L_λ chain mRNA sequences in MOPC-104E (Kyoto) myeloma were quantitated by hybridization with respective chain-type specific cDNA (Fig. 2). There is, as expected, extensive hybridization between RPC-20 cDNA and mRNA derived from a L_λ chain producing tumor, MOPC-104E (Kyoto). A $Crt_{1/2}$ value of 9×10^{-3} is comparable to RPC-20 mRNA (Fig. 1C). Unexpectedly, hybridization between MOPC-104E mRNA and MOPC-31C cDNA attained 60% hybridization with a $Crt_{1/2}$ value of 7×10^{-3}. Since the $Crt_{1/2}$ values are quite similar to each other, the amounts of L_κ and L_λ chain mRNA sequence in MOPC-104E mRNA are almost equivalent. The lower extent of hybridization indicates that the L_κ chain sequence present in MOPC-104E myeloma has a variable region sequence different from MOPC-31C L_κ chain. Similarly, almost equivalent amounts of L_κ and L_λ chain mRNA sequences were present in total cytoplasmic RNA prepared from MOPC-104E myeloma (NIH) which had been maintained independently from MOPC-104E (Kyoto)(Fig. 3). Polysomal RNA prepared from MOPC-104E (Osaka) also contained L_κ chain mRNA slightly more than L_λ chain mRNA (unpublished observation).

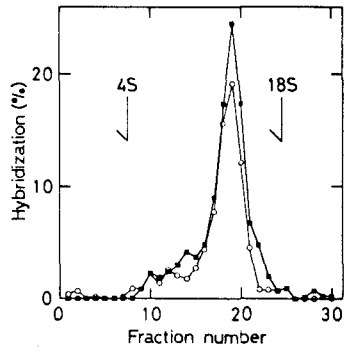


Fig. 4. Sucrose gradient centrifugation of MOPC-104E (NIH) RNA. After heat quenching, cytoplasmic RNA of MOPC-104E tumor (1.75mg) was layered on a 5-22% (w/v) sucrose gradient and centrifuged at 37,000rpm for 15h at 15°C in a Hitachi RPS 40T rotor (1,11). Fifteen µl of each fraction was hybridized with either MOPC-41 L_κ cDNA (1,270cpm) or RPC-20 L_λ cDNA (773cpm) in a 200µl reaction mixture for 30 min. Locations of 18S ribosomal RNA and tRNA (4S) markers were indicated in the figure. Hybridization with L_κ cDNA (O); L_λ cDNA (■). Endogenous S_1 nuclease resistance was subtracted from all values.

Characterization of light chain mRNA sequences present in MOPC-104E

The molecular size of L_κ and L_λ chain mRNA present in the MOPC-104E (NIH) cytoplasmic RNA fraction was analyzed by sedimentation in sucrose density gradient under conditions which abolish nonspecific aggregation of RNA (1)(Fig. 4). Both L_κ and L_λ chain mRNA sedimented together with a single peak of 12.5S which was known as the size of immunoglobulin light chain mRNA (1-8,11). The MOPC-104E L chain mRNA preparation purified by the method successfully employed for purification of MOPC-41 and RPC-20 mRNAs (1,5) migrated as a single band upon polyacrylamide gel electrophoresis in formamide, although this preparation contained equivalent amount of both L_κ and L_λ chain mRNA sequences. Both L_κ and L_λ chain mRNAs present in MOPC-104E contain poly(A) sequences because they were adsorbed to oligo(dT)-cellulose.

DISCUSSION

Available evidence indicates that three lines of MOPC-104E myeloma, which had been maintained in different laboratories, contain L_κ chain mRNA in a quantity almost equivalent to that of L_λ chain mRNA. Schechter (4) reported that the

λ type mRNA preparation purified from MOPC-104E myeloma contained about 3% of κ type mRNA. The discrepancy is probably attributable to their purification method which includes immunoprecipitation of polysomes using antibodies against L_λ chain protein. Myeloma tumors producing two classes of antibody have been observed (14). Investigators from several laboratories including our own (1,2,4) reported that light chain mRNA preparations from some myelomas contain a tiny amount of light chain mRNA sequences of the heterologous type. In these cases the major sequence is usually orders of magnitude more than the contaminant sequence, which could arise from circulating lymphocytes present in tumor tissues or from leaking transcription of the repressed gene in the myeloma cells.

Following questions should arise: 1) whether an individual MOPC-104E myeloma cell produces both L_κ and L_λ chain mRNA sequences; 2) whether the L_κ chain mRNA sequence is translated *in vivo* and the product is secreted from the cells. Hausman and Bosma (15) reported that they were unable to detect any L_κ chain protein in IgM secreted by MOPC-104E. However, a variant of MOPC-104E which was obtained by the alternate passage between *in vitro* culture and subcutaneous propagation in mice, produced IgM containing L_κ instead of L_λ chains and an altered μ type H chain. It is interesting to know whether the variant contains L_λ chain mRNA or not. At present we have no definitive answer to any of the above questions. We are trying to clone the MOPC-104E cell line and to study immunological properties of protein secreted by the myeloma, which will answer these questions.

ACKOWLEDGEMENT: This work was supported by grants from Jane Coffin Child Memorial Fund for Medical Research, from Matsunaga Science Foundation and from Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Honjo, T., Packman, S., Swan, D., Nau, M., and Leder, P. (1974) Proc. Nat. Acad. Sci. U.S., 71, 3659-3663.
2. Stavnezer, J., Huang, R.C.C., Stavnezer, E., and Bishop, J.M. (1974) J. Mol. Biol., 88, 43-63.
3. Rabbitts, T.H., and Milstein, C. (1975) Eur. J. Biochem., 52, 125-133.
4. Schechter, I. (1975) Proc. Nat. Acad. Sci. U.S., 72, 2511-2514.

5. Honjo, T., Swan, D., Nau, M., Norman, B., Packman, S., Polsky, F., and Leder, P. (1976) Biochemistry, 13, 2775-2779.
6. Honjo, T., Packman, S., Swan, D., and Leder, P. (1976) Biochemistry, 13, 2780-2784.
7. Faust C.H., Diggelman, H., and Mach, B. (1974) Proc. Nat. Acad. Sci. U.S., 71, 2491-2495.
8. Tonegawa, S. (1976) Proc. Nat. Acad. Sci. U.S., 73, 203-207.
9. McIntire, K.R., Asofsky, R.M., Potter, M., and Kuff, E.L. (1965) Science, 150, 361-362.
10. Namba, Y., and Hanaoka, M. (1972) J. Immunol., 109, 1193-1200.
11. Ono, M., Kondo, T., and Kawakami, M. (1976) J. Biochem. in the press
12. Ono, M., and Kawakami, M. (1976) Jap. J. Microbiol., 20, 203-208.
13. Kirby, K.S. (1968) Methods Enzymol., 12B, 87-99.
14. Wang, A.C., Wang, I.Y.F., McCormick, N.J., and Fudenberg, H.H. (1969) Immunochemistry, 6, 451-459.
15. Hausman, S.J., and Bosma, M.J. (1975) J. Exp. Med., 142, 998-1010.

[Reprinted from Biochemistry, (1977) 16, 5764.]

Accumulation of Immunoglobulin Messenger Ribonucleic Acid in Immunized Mouse Spleen†

Tasuku Honjo,* Akira Shimizu, Masa-aki Tsuda, Shunji Natori, Tohru Kataoka, Chieko Dohmoto, and Yoshitake Mano

ABSTRACT: We have measured the concentration of mRNAs coding for immunoglobulins, κ and λ type light chains and γ_1 type heavy chain, in mouse spleen cells activated by bacterial lipopolysaccharide or sheep red blood cells. These mRNAs were quantitated by hybridization to radioactive DNA complementary to highly purified immunoglobulin mRNAs from mouse myelomas. In the lipopolysaccharide-stimulated spleen cells, only light chain mRNA accumulated, whereas γ_1 type heavy chain mRNA remained unvaried. The light chain mRNA concentration also increased in purified bone-marrow-derived lymphocytes. The lipopolysaccharide-induced light chain mRNA was similar to light chain mRNAs purified from myelomas. The accumulation and disappearance of light chain mRNA in bone-marrow-derived lymphocytes coincide with the kinetics of synthesis of immunoglobulin M which is the major species induced by lipopolysaccharide. In sheep red blood cell stimulated spleen, the specific accumulation of κ type light chain and γ_1 type heavy chain mRNAs parallels immunoglobulin G synthesis. These results seem to indicate that the increment of immunoglobulin mRNA concentration in bone-marrow-derived lymphocytes is important for induction of immunoglobulin synthesis.

In order to elucidate the molecular mechanisms of the immune response, it is essential to know which step(s) of antibody synthesis is stimulated upon activation of immunocompetent lymphocytes. A vital question is whether induction of antibody synthesis is due to activation of Ig[1] gene transcription or of translational steps of preexisting Ig mRNA. Purification of

† From Department of Physiological Chemistry and Nutrition, Faculty of Medicine and Faculty of Pharmaceutical Sciences, University of Tokyo, Hongo, Tokyo, Japan. *Received April 25, 1977; revised manuscript received September 1, 1977.* This investigation was supported by grants from Jane Coffin Child Memorial Fund for Medical Research, from Matsunaga Science Foundation, and from Ministry of Education, Science and Culture of Japan.

[1] Abbreviations used are: Ig, immunoglobulin; L_κ and L_λ chains, κ and λ type light chains; H_{γ_1} chain, γ_1 type heavy chain; LPS, lipopolysaccharide; SRBC, sheep red blood cells; C_rt, the product of the concentration of RNA and the time of incubation; T and B lymphocytes, thymus-derived and bone-marrow-derived lymphocytes; poly(A), poly(adenylic acid).

Ig L and H chain mRNAs and their reverse transcription have provided highly radioactive DNA complementary to the Ig mRNA for a nucleic acid hybridization probe to quantitate mRNA (Honjo et al., 1974, 1976a,b; Faust et al., 1974; Stavnezer et al., 1974; Rabbitts, 1974; Schechter, 1975; Ono et al., 1977a,b). Using this probe, we have measured the Ig mRNA content in mouse spleen cells which had been immunized with bacterial LPS or SRBC.

LPS which is commonly termed B lymphocyte mitogen activates polyclonal B lymphocytes to induce IgM synthesis in the absence of any accessory cells such as T lymphocytes or macrophages (Coutinho and Möller, 1975). In contrast a normal antigen like SRBC requires cooperation of the accessory cells to induce IgM and IgG syntheses in B lymphocytes.

We will report here that Ig mRNA accumulates and disappears specifically coincidental to Ig synthesis in the LPS (or SRBC)-stimulated spleen. Accumulation of Ig mRNA is due to the increased concentration of mRNA in a B lymphocyte but not due to the increased number of B lymphocytes containing a constant amount of Ig mRNA.

Experimental Procedures

(a) *Animals and Immunization.* Six to 7 week old dd/Y strain mice received intraperitoneal injection of 250 μg/0.2 mL LPS. SRBC (2×10^8 cells/0.2 mL) were injected from the tail vein. On days indicated five mice were sacrificed, and spleens were removed. Spleens were immediately frozen and kept at −80 °C until RNA extraction. LPS was prepared from *Escherichia coli* according to Westphal et al. (1952). SRBC was obtained commercially.

(b) *Preparation of RNA.* MOPC 41 L_κ chain mRNA was purified from myeloma as described before (Honjo et al., 1974). MOPC 31C L_κ chain mRNA was a generous gift of Dr. Ono (Ono et al., 1977a). To quantitate mRNA in spleen cells, RNA was extracted from cytoplasmic fraction of spleen by a slight modification of the method described by Kirby (1968). Spleens were homogenized with a Potter-type homogenizer (9 strokes) in 10 volumes of 0.25 M sucrose containing 50 mM Tris-HCl (pH 9.0), 25 mM KCl, 5 mM $Mg(CH_3COO)_2$, 0.5 mM cycloheximide, and 0.1% diethyl pyrocarbonate. The homogenate was centrifuged for 10 min at 5000*g* at 0 °C. To the supernatant was added sodium dodecyl sulfate to 0.3%. The mixture was extracted with an equal volume of phenol–*m*-cresol mixture (water-saturated phenol/*m*-cresol, 9:1, v/v) containing 0.1% 8-hydroxyquinoline. After centrifugation the aqueous phase was removed and extraction was repeated until no more white material was visible at the interphase. The RNA was precipitated with ethanol at −20 °C. The precipitated RNA was dissolved in H_2O and stored at −80 °C.

(c) *Preparation of cDNA.* L_κ chain cDNA was prepared with avian myeloblastosis virus reverse transcriptase (a gift from Dr. Leder) from L_κ chain mRNA purified from MOPC 41 myeloma (Honjo et al., 1974). L_λ chain cDNA was prepared from RPC 20 myeloma mRNA (Honjo et al., 1976a,b). Preparation of H_{γ_1} chain cDNA from MOPC 31C H_{γ_1} chain mRNA was described previously (Ono et al., 1977b). Globin cDNA was synthesized according to Ross et al. (1973). The specific radioactivity of cDNAs was 1×10^7 cpm/μg. L chain cDNAs, H_{γ_1} chain cDNA, and globin cDNA thus prepared were approximately 490, 600, and 400 nucleotides long, respectively, as measured by polyacrylamide gel electrophoresis in formamide (Honjo et al., 1976a). Most of the sequences of the Ig cDNAs employed in the present study are complementary to the constant region sequences which are common to all the Ig mRNAs of the respective subclasses (Honjo et al., 1974, 1976a,b; Ono et al., 1977b). Accordingly, hybridization to these cDNAs allows us to determine the total concentration of the Ig mRNAs which have a variety of variable region sequences.

(d) *Quantitation of mRNAs in Spleen RNA.* Increasing amounts of splenic RNA were hybridized to respective [^{3}H]cDNA in 0.6 M NaCl at 75 °C. Hybrids formed were assayed by S1 nuclease digestion as described before (Honjo et al., 1974). The initial hybridization rate (percentage hybridized per C_rt) was determined from the slope obtained as shown in Figure 1 (Marbaix et al., 1975; Orkin et al., 1975). Under the conditions employed the amount of hybridizable RNA is at least in fivefold excess over that of hybridized cDNA. The fraction of mRNA in total RNA is calculated from the ratio of the initial hybridization rate of spleen RNA to that of purified mRNA, assuming that the purified mRNA is 100% pure. The initial hybridization rates (hybridization percentage per C_rt) of MOPC 41 mRNA (L_κ), RPC 20 (L_λ), MOPC 31C mRNA (H_{γ_1}), and mouse globin mRNA are 1.66×10^5, 1.66×10^5, 1.38×10^5, and 2.4×10^5, respectively, under the same conditions. The average number of mRNA molecules in a spleen cell was calculated as follows: (total RNA content per cell) × (fraction of mRNA in total RNA) × 6.02×10^{23}/(molecular weight of mRNA). L_κ chain mRNA, L_λ chain mRNA, H_{γ_1} chain mRNA, and globin mRNA (α and β chains combined) are 4.5×10^5, 4.0×10^5, 7.0×10^5, and 4.2×10^5 daltons, respectively (Honjo et al., 1974, 1976a; Ono et al., 1977b).

(e) *Purification of B Lymphocytes from Mouse Spleen.* Spleen cells were suspended in Gey's solution (Gey and Gey, 1936) containing 0.08 M NH_4Cl for 2 min at 0 °C to lyse erythrocytes and erythroid cells. The residual cells were spun down and suspended in Eagle's minimal essential medium to make 1×10^7 spleen cells/mL. T lymphocytes were lysed by rabbit antiserum directed against mouse brain-associated T lymphocyte antigen (a gift from Dr. Okumura of Chiba University) and guinea pig complement by the method of Sato et al. (1976). B lymphocytes which survived the anti-T-lymphocyte serum treatment were spun down, washed several times by centrifugation, and frozen at −80 °C until RNA extraction. Alternatively, B lymphocytes were isolated by adsorption to nylon wool (Julius et al., 1973).

(f) *Determination of IgM- and IgG-Producing Cell Number.* LPS induces production of IgM against a variety of antigens. We have counted the number of spleen cells which produce IgM against SRBC as a measure of IgM synthesis in spleen. The indirect plaque assay was used to measure the sum of IgG- and IgM-producing cells (Šterzl and Říha, 1965; Dresser and Wortis, 1965). The number of IgG-producing cells was determined as the difference of the direct plaque count (IgM-producing cells) from the indirect plaque count. The number of plaque-forming cells was determined according to Cunningham (1965). The plaque numbers were counted in four slides which contain 10^5 to 10^6 cells for each time point and the average numbers were plotted.

(g) *Other Methods.* Thermal stabilities of the hybrids formed between spleen RNA and cDNA were tested by S1 nuclease digestion as described by Honjo et al. (1976b). The size of RNA was determined by centrifugation in a sucrose gradient (5 to 22%) as described before (Honjo et al., 1974). RNA was heated at 70 °C for 10 min and immediately chilled before centrifugation. DNA and RNA were determined colorimetrically according to Schneider (1957).

Results

(a) *Quantitation of L Chain mRNA Sequence in Spleen*

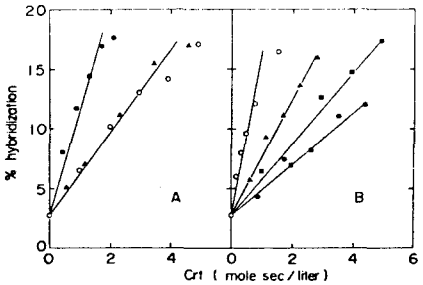


FIGURE 1: Hybridization of LPS-stimulated spleen RNA to L_κ chain cDNA. Increasing amounts of splenic RNA were hybridized to a constant amount of L_κ chain [^{3}H]cDNA as described in Experimental Procedures. Each reaction mixture (250 μL) contained 800 cpm of cDNA and was incubated for 30 min. Fractions of L_κ chain mRNA were calculated from slopes as described in Experimental Procedures and summarized in Table I. Symbols represents spleen RNA obtained on various days after LPS injection. (A) (○) Day 0; (▲) day 1; (●) day 2; (B) (○) day 3; (▲) day 5; (●) day 7; (■) day 10.

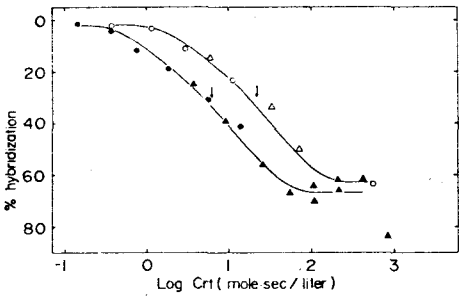


FIGURE 2: Hybridization kinetics of spleen RNA to L_κ chain cDNA. RNAs obtained from normal spleen and from LPS-stimulated spleen (day 3) were hybridized to L_κ chain cDNA. Aliquots were removed at time intervals and hybrids formed were assayed as described in Experimental Procedures. Spleen RNA used were the same batch as used in Figure 1. (○) Normal spleen RNA, 2.2 mg/mL; (△) 3.5 mg/mL of normal spleen RNA; (●) 0.56 mg/mL of LPS-stimulated spleen RNA; (▲) 5.4 mg/mL of LPS-stimulated spleen RNA.

RNA. Cytoplasmic RNA was extracted from mouse spleen at various time intervals after LPS injection and tested for its hybridizability to [^{3}H]cDNA complementary to the L_κ chain mRNA purified from a mouse myeloma, MOPC 41. Our assay scores all L_κ chain mRNA since RNA was hybridized to cDNA, 78% of whose sequence is complementary to the constant region sequence (see below). Figure 1 shows the initial rates of hybridization obtained from the plot of the extent of hybridization vs. C_rt. The sequence hybridizable to L_κ chain cDNA began to increase 2 days after LPS injection, reaching a maximum (4.4-fold) on day 3. The hybridizable sequence decreased rather quickly, going back to the normal level by day 7. The results indicate that the L_κ chain mRNA sequence accumulated temporarily in mouse spleen by LPS injection. The concentration of L_κ chain mRNA sequence increased similarly in total cellular RNA extracted from LPS-treated spleen.

(b) *Characterization of L_κ Chain mRNA Sequence in Spleen.* Hybridization kinetics between L_κ chain cDNA and RNA derived from day 0 and day 3 spleens gave $C_rt_{1/2}$ values of 26 and 6.3, respectively, as shown in Figure 2. Since the $C_rt_{1/2}$ value is inversely proportional to the concentration of the sequence, the ratio of the $C_rt_{1/2}$ values indicates that 4.1-fold greater concentration of the L_κ chain mRNA sequence is present in day 3 spleen RNA than in day 0 spleen RNA, which is in good agreement with the results measured by the initial rate of hybridization as shown above. The maximal extent of hybridization reached 64% to 67% with RNAs from day 0 and day 3 spleens. The extents of hybridization likely

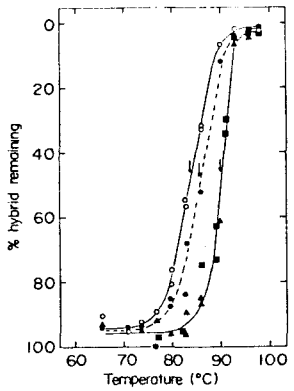


FIGURE 3: Thermal stability of hybrids formed between L_κ chain cDNA and spleen RNA. L_κ chain cDNA was hybridized with normal spleen RNA, LPS-stimulated spleen RNA (day 3), and L_κ chain mRNAs purified from MOPC 41 and MOPC 31C myelomas until C_rt values of 470, 450, and 0.2, respectively, were attained. The hybrids formed were tested for their thermal stability as described in Experimental Procedures. T_m values for hybrids with normal spleen RNA, LPS-stimulated spleen RNA, MOPC 41 mRNA, and MOPC 31C mRNA were 84, 86, 90, and 90 °C, respectively. (○) Normal spleen RNA; (●) LPS-stimulated spleen RNA; (▲) MOPC 41 mRNA; (■) MOPC 31C mRNA.

reflect that proportion of the cDNA complementary to the L_κ chain constant region sequence (Honjo et al., 1974). Hybridization with L_κ chain mRNAs purified from myelomas, MOPC 41 (homologous) and MOPC 31C (heterologous), attained the maximal extent of 73% and 65% (data not shown), respectively, indicating that 78% of the hybridizable sequence of the cDNA preparation employed represents the constant region sequence. Figure 3 shows the thermal stability of the hybrids. The melting profiles of the hybrids were reasonably sharp and T_m values for the hybrids with normal spleen RNA, LPS-stimulated spleen RNA, and purified L_κ chain mRNAs (both MOPC 41 and MOPC 31C mRNAs) were 84, 86, and 90 °C, respectively. The results indicate that most of the sequences in the spleen RNA hybrids are congruently base paired. Nonetheless, a slight reduction of the T_m values for the splenic RNAs compared with those of the purified L_κ chain mRNAs could be due to partially inaccurate base pairing in the variable region and/or untranslated sequences which constitute approximately 20% of the cDNA sequence (Honjo et al., 1974).

When the RNA fractions from normal as well as LPS-stimulated spleen were treated with RNase or alkali, hybridization with the L_κ chain cDNA was completely abolished, indicating that the sequence assayed is present in RNA molecules. Heat denaturation of normal spleen RNA did not increase the amount of the sequence hybridizable to L_κ chain cDNA. This excludes the possibility that a double-stranded "precursor" RNA exists in normal spleen and is converted to a hybridizable form by the LPS stimulation. The sequence hybridizable to the L_κ chain cDNA from both normal and LPS-stimulated spleens sedimented with a major peak at 13 S upon centrifugation in sucrose gradient under the conditions which abolish nonspecific aggregation of RNA (Honjo et al., 1974), although some degradation of RNA was inevitable. The sedimentation profile of myeloma L chain mRNA shows a single peak at 13–14 S (Honjo et al., 1976a). The L_κ chain mRNA sequences in splenic RNA were adsorbed to oligo(dT)-cellulose, suggesting that they contain poly(A) sequences.

(c) *L_κ Chain mRNA Accumulation and IgM Synthesis.* Intraperitoneal injection of LPS brings about a variety of biochemical as well as cytological changes in mouse spleen

TABLE I: Quantitation of L_κ Chain mRNA in LPS-Stimulated Spleen.

Days after LPS injection	(Cells[a]/ spleen) $\times 10^{-8}$	RNA/ cell (pg)	RNA/ spleen (mg)	Initial rate of hybridization[b]	(mRNA/ total RNA) $\times 10^5$	mRNA molecules/ cell	(mRNA molecules/ spleen) $\times 10^{-10}$
0	3.1	2.11	0.65	3.15	1.90	53	1.65
1	2.5	4.19	1.06	3.15	1.90	105	2.63
2	3.6	3.98	1.44	9.20	5.54	292	10.5
3	5.8	3.05	1.76	14.0	8.43	340	19.7
5	4.7	3.07	1.45	4.85	2.92	119	5.58
7	7.9	2.89	2.28	2.15	1.30	50	3.93
10	5.6	2.37	1.39	3.00	1.81	57	3.18

[a] Cell numbers were calculated from DNA contents assuming that a mouse spleen cell contains 6 pg DNA. [b] Initial rates of hybridization were expressed as percentage hybridized per C_rt which was calculated from Figure 1.

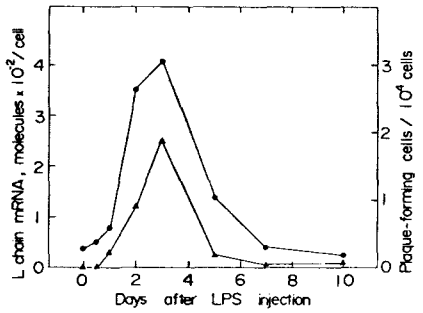


FIGURE 4: Effects of LPS injection on spleen cells. IgM-producing cell numbers were determined as described in Experimental Procedures. The number of L_κ chain mRNA molecules per spleen cell were determined as described in Figure 1. (●) Average number of L_κ chain mRNA molecules in a spleen cell; (▲) number of plaque-forming cells against SRBC.

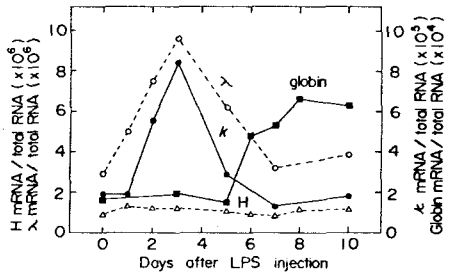


FIGURE 5: Kinetics of Ig mRNA accumulation in LPS-stimulated spleen. Fraction of each mRNA was determined by hybridization to a constant amount of respective cDNA exactly as described in Figure 1 and Experimental Procedures. Data for L_κ chain mRNA were taken from Table II. Spleen RNA used was the same batch as used in Figure 1 and Table II. (●) L_κ chain mRNA; (○) L_λ chain mRNA; (△) $H_{\gamma 1}$ chain mRNA; (■) globin mRNA.

(Gronowicz and Coutinho, 1974; Takano et al., 1967; Fruhman, 1966). Consistent with previous reports from other laboratories (Coutinho and Möller, 1975) that both DNA and RNA syntheses are stimulated by LPS injection, the number of cells per spleen almost doubled on day 3 and the amount of RNA per spleen increased by 2.7-fold as shown in Table I. Assuming that the purified L_κ chain mRNA is 100% pure we can calculate the average number of L_κ chain mRNA molecules present in a spleen cell or in a spleen. On day 3 the fraction of L_κ chain mRNA, the average number of L_κ chain mRNA molecules per cell, and the average number of L_κ chain mRNA molecules per spleen were elevated by 4.4-fold, 6.4-fold, and 11.9-fold, respectively. IgM synthesis as measured by the number of plaque-forming cells against SRBC increased and reached a maximal level on day 3, followed by a rapid disappearance by day 7 as shown in Figure 4. It is clear that the accumulation of L_κ chain mRNA parallels the increase of IgM synthesis.

(d) *Specific Accumulation of L Chain mRNA.* In order to see whether the accumulation of L_κ chain mRNA is specific we have measured the concentration of several other mRNAs in LPS-stimulated spleen. Figure 5 shows such studies on L_λ chain mRNA, $H_{\gamma 1}$ chain mRNA and globin mRNA. The fractional content of L_λ chain mRNA, as expected, increased by 3.3-fold on day 3 though the concentration of L_λ chain mRNA is one order of magnitude less than that of L_κ chain mRNA. L_λ chain mRNA seems to accumulate and disappear in good coincidence with L_κ chain mRNA. On the other hand, the concentration of $H_{\gamma 1}$ chain mRNA remained essentially constant. The results are consistent with the reports from other laboratories that a single injection of LPS does not induce IgG synthesis in vivo (Britton and Möller, 1968; Coutinho and Möller, 1975). The globin mRNA concentration remained constant until day 5. On day 6, however, a significant increase (3-fold) of the mRNA concentration was seen which was conserved until day 10. The results confirm the report that LPS stimulates the erythropoietic activity in mouse spleen (Fruhman, 1966).

TABLE II: Accumulation of L_κ Chain mRNA in LPS-Stimulated B Lymphocytes.

Method of B-lymphocyte isolation[a]	Fraction of L_κ chain mRNA ($\times 10^5$)	
	Control	LPS (day 3)
Nylon wool	4.92	17.5
anti-T-lymphocyte serum	1.02	6.50

[a] Methods of B-lymphocyte isolation were described in Experimental Procedures.

(e) *L_κ Chain mRNA Content in Purified B Lymphocytes.* We have measured the L_κ chain mRNA concentration in B lymphocytes which were purified from normal and LPS-stimulated spleen cells by adsorption to nylon wool or by the treatment with rabbit antiserum directed against mouse T lymphocytes. Table II shows that the L_κ chain mRNA concentration increased by 3.6- to 6.3-fold in B lymphocytes derived from LPS-stimulated spleen. The L_κ chain mRNA concentrations in normal and LPS-stimulated B lymphocytes were comparable to those in normal and LPS-stimulated spleen cells, respectively. Since about 50% of spleen cells are B lymphocytes, they seem to contain a major portion, if not all, of L_κ chain mRNA in spleen.

TABLE III: Quantitation of L_κ Chain mRNA in SRBC-Stimulated Spleen.

Days after SRBC injection	(Cells[a]/ spleen) $\times 10^{-8}$	RNA/ cell (pg)	RNA/ spleen (mg)	Initial rate of hybridization[b]	(mRNA/ total RNA) $\times 10^5$	mRNA molecules/ cell	(mRNA molecules/ spleen) $\times 10^{-10}$
0	4.3	1.58	0.679	2.04	1.23	26	1.11
1	4.0	2.05	0.820	2.78	1.67	48	1.91
2	4.0	1.70	0.680	2.70	1.63	37	1.47
3	4.0	2.71	1.08	2.88	1.73	62	2.48
5	4.0	2.35	0.940	3.53	2.13	66	2.65
7	3.8	1.89	0.718	5.23	3.15	79	2.99
10	4.1	1.88	0.771	5.97	3.60	90	3.67

[a] Cell numbers were calculated from DNA contents assuming that a mouse cell contains 6 pg of DNA. [b] Initial hybridization rates were expressed as percentage hybridized per C_rt which was determined as described in Figure 1.

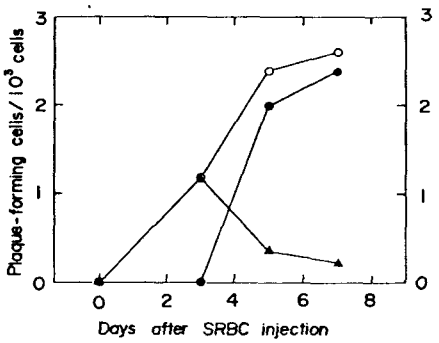


FIGURE 6: Increase of IgM- and IgG-producing cells in SRBC-stimulated spleen. Numbers of IgM- and IgG-producing cells were determined as described in Experimental Procedures at time intervals indicated after SRBC injection. (○) Total number of plaque-forming cells; (▲) number of IgM-producing cells; (●) number of IgG-producing cells.

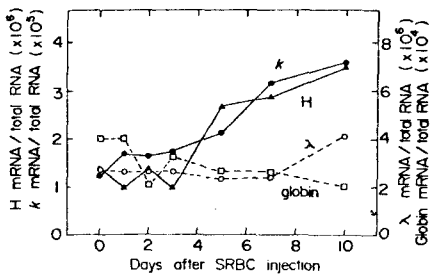


FIGURE 7: Kinetics of Ig mRNA accumulation in SRBC-stimulated spleen. Fraction of each mRNA was determined by hybridization to a constant amount of respective cDNA exactly as described in Figure 1 and Experimental Procedures. The same batch of spleen RNA was used throughout the experiments. (●) L_κ chain mRNA; (○) L_λ chain mRNA; (▲) H_{γ_1} chain mRNA; (□) globin mRNA.

(f) *Ig mRNA Accumulation in SRBC-Stimulated Spleen.* It is important to see whether a normal antigen also induces accumulation of Ig mRNA or not. We have used SRBC which is known as a potent antigen. As shown in Figure 6, SRBC stimulates IgM synthesis in mouse spleen within 2–3 days after its injection, followed by extensive IgG synthesis around day 5 to 7 in confirmation of the results reported by Dresser and Wortis (1965) and Šterzl and Říha (1965). Table III shows the kinetics of the spleen cell number, RNA content, and the L_κ chain mRNA concentration in SRBC-stimulated spleen. In contrast to LPS stimulation, SRBC injection did not increase the number of spleen cells. The RNA content was slightly elevated (1.6-fold per spleen at most). The fractional mRNA content rose slightly on day 1 and then increased significantly on day 5, reaching a 2.9-fold higher level on day 10. Similar studies on L_λ chain mRNA, H_{γ_1} chain mRNA, and globin mRNA are shown in Figure 7. The concentration of L_λ chain mRNA maintained the original level until day 7 and increased slightly (1.5-fold) on day 10. The H_{γ_1} chain mRNA concentration began to increase extensively on day 5, reaching a 2.5-fold higher level on day 10. The results are in agreement with the results that SRBC injection induces IgG synthesis in a later stage (days 5–7) of the immune response (Figure 6). There was no increment of the globin mRNA content throughout 10 days after SRBC injection. These results show that an ordinary antigen like SRBC also induces the specific accumulation of Ig mRNAs concurrently to the increase of IgG synthesis.

Discussion

The final step of the immune response, which is manifested as induction of Ig synthesis in B lymphocytes, might be controlled in terms of amplification of the Ig genes or in terms of inducing their transcription or translation of the corresponding mRNAs. The genes coding for L chain mRNA have been shown to be repeated not more than five times in the mouse genome regardless of whether the cells from which DNA was obtained are actively engaged in Ig synthesis or not (Honjo et al., 1974, 1976b; Stavnezer et al., 1974; Faust et al., 1974). The results rule out the possibility that gene amplification plays an essential role for the immune response.

We have shown accumulation of Ig mRNAs in mouse spleen cells as they divide and differentiate into high-rate antibody-secreting lymphocytes in response to LPS or SRBC injection. Since the concentration of L_κ chain mRNA increases in the total cellular RNA including nuclear as well as cytoplasmic RNA, it is unlikely that the accumulation of Ig mRNA is attributable to the increase of nucleocytoplasmic transport. The apparent accumulation of Ig mRNAs in spleen could be also due to an increase of the B-lymphocyte population in a spleen while each B lymphocyte contains the unvaried amounts of Ig mRNAs. The above possibility was excluded by the experiment in which purified B lymphocytes were shown to contain an elevated amount of L_κ chain mRNA per cell upon LPS stimulation (Table II). Furthermore, the quantitative consideration makes the above possibility unlikely. Shands et al. (1973) have reported that about 50% of spleen cells are activated by LPS stimulation. Since approximately 50% of spleen cells are B lymphocytes in normal spleen (Raff et al., 1971; Storb et al., 1976), LPS seems to stimulate a majority of B lymphocytes in spleen. Therefore, the enlargement of the B-lymphocyte population within a spleen cannot increase the average Ig mRNA content per spleen cell more than twice as much as the normal level.

Since the time course of the accumulation and disappear-

ance of the L chain mRNA agrees with kinetics of IgM synthesis, the number of mRNA molecules present seems to be closely related with the Ig synthetic capacity of lymphocytes. This implies that regulation involving the translational activation of a reserve population of L chain mRNA does not seem to play a major role in induction of Ig synthesis. Although our data do not allow us to conclude whether the accumulation of L chain mRNA is attributable to the increase of the Ig gene transcription or to the reduction of mRNA degradation, preliminary studies which have been undertaken in in vitro cultured spleen cells show that transcription of the L_κ chain gene is enhanced by LPS stimulation (Tsuda, Natori, and Honjo, manuscript in preparation).

It is apparent that there are quantitative as well as qualitative similarities between L_κ chain mRNA from spleen and from mouse myelomas. Both mRNAs have poly(A) sequences and a similar size. The extent of hybridization of splenic RNA to MOPC 41 cDNA is similar to that of heterologous L_κ chain mRNA purified from MOPC 31C myeloma. The hybrids with splenic mRNA have the T_m values slightly lower than myeloma mRNA hybrids. The difference of the T_m values (4 to 6 °C) corresponds to a maximum of 4 to 6% mismatching (Britten et al., 1974). It is plausible that partially mismatched hybrids are formed between MOPC 41 cDNA and splenic mRNA, which is a mixture of hundreds of L_κ chain mRNA species, each containing different variable region sequences. In fact, the extent of hybridization of splenic mRNA is similar to, nonetheless, usually slightly higher than that of the heterologous L_κ chain mRNA. The decreased thermal stability of the hybrids could be ascribed to several other factors such as length of hybridizing RNA and/or presence of L chain-like RNA of unknown function. Storb et al. (1976) reported that the L_κ chain cDNA hybrid with spleen RNA melts 10.7 °C below the T_m value of the hybrid with homologous mRNA.

The kinetics of Ig mRNA accumulation agree with those immunological observations which have been known to be induced by LPS or SRBC injection. LPS injection induces accumulation of L chain mRNA but not of $H_{\gamma 1}$ chain mRNA, corresponding to stimulation of IgM synthesis but not of IgG synthesis. SRBC-induced accumulation of L chain mRNA and $H_{\gamma 1}$ chain mRNA coincides with IgG synthesis. It was rather unexpected that SRBC injection did not increase L chain mRNAs significantly at an early stage (days 2-3) when spleen cells produce IgM. This is probably due to the fact that SRBC injection induces much smaller quantities of IgM synthesis than does LPS injection. The increase of the globin mRNA level coincides with the report that LPS injection stimulates erythropoiesis in a week or so in mouse spleen (Fruhman, 1966). Our preliminary data show that LPS does not stimulate the erythropoietic activity in cultured spleen cells (Tsuda, Natori, and Honjo, manuscript in preparation), suggesting that LPS-induced erythropoiesis in spleen is a secondary effect of the bacterial endotoxin.

Investigators from another laboratory reported the quantitation of Ig H chain mRNA in LPS-stimulated mouse spleen cells cultured in vitro (Stevens et al., 1975). The assay method employed for mRNA measurement, however, was shown to be difficult to reproduce (Stevens and Williamson, 1975).

Acknowledgments

We thank Dr. D. Mizuno for his support and encouragement and Dr. P. Leder for a generous gift of reverse transcriptase. The manuscript was improved by the critical reading by Drs. P. Leder and R. Crouch. Thanks are also extended to Drs. K. Okumura and T. Tada for helpful advice and suggestions.

References

Britten, R. J., Graham, D. E., and Neufeld, B. R. (1974), *Methods Enzymol. 29,* 363-418.

Britton, S., and Möller, G. (1968), *J. Immunol. 100,* 1326-1334.

Coutinho, A., and Möller, G. (1975), *Adv. Immunol. 21,* 113-238.

Cunningham, A. J. (1965), *Nature (London) 207,* 1106-1107.

Dresser, D. W., and Wortis, H. H. (1965), *Nature (London) 208,* 859-861.

Faust, C. H., Diggelmann, H., and Mach, B. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 2491-2495.

Fruhman, G. J. (1966), *Blood 27,* 363-370.

Gey, G. O., and Gey, M. K. (1936), *Am. J. Cancer 27,* 45-76.

Gronowicz, E., and Coutinho, A. (1974), *Eur. J. Immunol. 4,* 771-776.

Honjo, T., Packman, S., Swan, D., and Leder, P. (1976b), *Biochemistry 15,* 2780-2785.

Honjo, T., Packman, S., Swan, D., Nau, M., and Leder, P. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 3659-3663.

Honjo, T., Swan, D., Nau, M., Norman, B., Packman, S., Polsky, F., and Leder, P. (1976a), *Biochemistry 15,* 2775-2779.

Julius, M. H., Simpson, E., and Herzenberg, L. A. (1973), *Eur. J. Immunol. 3,* 645-649.

Kirby, K. S. (1968), *Methods Enzymol. 12B,* 87-99.

Marbaix, G., Huez, G., Burny, A., Cleuter, Y., Hubert, E., Leclercq, M., Chantrenne, H., Soreq, H., Nudel, U., and Littauer, U. Z. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 3065-3067.

Ono, M., Kondo, T., and Kawakami, M. (1977a), *J. Biochem. (Tokyo) 81,* 941-947.

Ono, M., Kondo, T., Kawakami, M., and Honjo, T. (1977b), *J. Biochem. (Tokyo) 81,* 949-954.

Orkin, S. H., Swan, D., and Leder, P. (1975), *J. Biol. Chem. 250,* 8753-8760.

Rabbitts, T. H. (1974), *FEBS Lett. 42,* 323-326.

Raff, M. C., Nase, S., and Mitchison, N. A. (1971), *Nature (London) 230,* 50-51.

Ross, J., Aviv, H., and Leder, P. (1973), *Arch. Biochem. Biophys. 158,* 494-502.

Sato, V. L., Waksal, S. D., and Herzenberg, L. A. (1976), *Cell. Immunol. 24,* 173-185.

Schechter, I. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 2511-2514.

Schneider, W. C. (1957), *Methods Enzymol. 3,* 680-684.

Shands, J. W., Jr., Peavy, D. L., and Smith, R. T. (1973), *Am. J. Pathol. 70,* 1-24.

Stavnezer, J., Huang, R. C., Stavnezer, E., and Bishop, J. M. (1974), *J. Mol. Biol. 88,* 43-63.

Šterzl, J., and Říha, I. (1965), *Nature (London) 208,* 858-859.

Stevens, R. H., Askonas, B. A., and Welstead, J. L. (1975), *Eur. J. Immunol. 5,* 47-54.

Stevens, R. H., and Williamson, A. R. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 4679.

Storb, U., Hager, L., Putnam, D., Buck, L., Fabin, F., and Clagett, J. (1976), *Proc. Natl. Acad. Sci. U.S.A. 73,* 2467-2471.

Takano, T., Yoshioka, O., and Mizuno, D. (1967), *Proc. Jpn. Soc. RES 7,* 13-18.

Westphal, O., Lüderitz, O., and Bister, F. (1952), *Z. Naturforsch. B 7,* 148-155.

PURIFICATION AND SOME PROPERTIES OF BOVINE PINEAL TRYPTOPHAN 5-MONOOXYGENASE.

Toshihiro Nukiwa*, Chiharu Tohyama**, Chieko Okita, Tohru Kataoka and Arata Ichiyama***

Department of Physiological Chemistry and Nutrition
Faculty of Medicine, University of Tokyo
Hongo, Bunkyo-ku, Tokyo 113, Japan

Received August 20,1974

SUMMARY : Tryptophan 5-monooxygenase was purified approximately 1,000-fold from the bovine pineal gland. The purified enzyme catalyzed the hydroxylations of both L-tryptophan and L-phenylalanine at a comparable rate. Evidence was presented suggesting that the hydroxylations of both amino acids were catalyzed by the single enzyme. The apparent Km values for L-tryptophan and for L-phenylalanine were approximately 16 and 32 μM, respectively, when tetrahydrobiopterin was used as a cofactor. The apparent molecular weight of the enzyme was estimated to be approximately 30,000 by gel filtration on columns of Sephadex G-75 and G-100 and by ultracentrifugation in sucrose density gradients. These properties of bovine pineal tryptophan 5-monooxygenase were distinguishable from those of rat liver phenylalanine hydroxylase, another enzyme which had been shown to catalyze the hydroxylations of both L-tryptophan and L-phenylalanine.

Tryptophan 5-monooxygenase (E.C.1.99.1.4) catalyzes the hydroxylation of L-tryptophan to L-5HTP[+], the initial reaction in the biosynthesis of serotonin in the brain and of melatonin in the pineal gland. The enzyme activity has been detected in various cell-free preparations from these organs (1-7) but the enzyme has not been extensively purified mainly due to low activity in tissue extracts and poor yields upon attempts at purification.

In a previous report from this laboratory (6), tryptophan 5-monooxygenase of the bovine pineal gland was shown to exist largely as an inactive form in tissue extracts. Preincubation of the inactive enzyme under anaerobic conditions in the presence of DTT resulted in an approximately 100-fold activation. During the course of this study, we have observed that the inactive form of the

---

* Present address, Department of Medical Chemistry, Faculty of Medicine Kyoto University, Yoshida, Sakyo-ku, Kyoto 606, Japan.

** On leave from the Department of Public Health, Faculty of Medicine, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan.

*** Present address, Department of Biochemistry, Hamamatsu University School of Medicine, Hamamatsu-shi, Shizuoka 432, Japan.

+ Abbreviations used are : 5HTP, 5-hydroxytryptophan; BpH_4, 5,6,7,8-tetrahydrobiopterin; $DMPH_4$, 6,7-dimethyl-5,6,7,8-tetrahydropterin; $6MPH_4$, 6-methyl-5,6,7,8-tetrahydropterin; DTT, dithiothreitol; BSA, bovine serum albumin; Tryptophan-1-^{14}C, tryptophan-(side chain-1)^{14}C; Tyrosine-1-^{14}C, tyrosine-(side chain-1)^{14}C.

enzyme was much more stable than the activated enzyme. In the present study, therefore, tryptophan 5-monooxygenase was purified as the inactive form about 1,000-fold from the bovine pineal gland. The enzyme activity was determined each time after the enzyme was activated. This communication describes the procedure for the purification and some properties of the purified enzyme.

MATERIALS AND METHODS

Bovine pineal glands were obtained at a slaughterhouse and stored frozen at -80° until use. L-tryptophan-1-^{14}C and L-tyrosine-1-^{14}C (New England Nuclear), L-phenylalanine-4-T (Radiochemical Centre), $DMPH_4$, $6MPH_4$ and DTT (Calbiochem), hydroxylapatite (Clarkson), P-cellulose and DEAE-cellulose (Brown), Sephadex G-75, G-100 and G-200 and CM-Sephadex C-50 (Pharmacia) were purchased from the sources indicated in parentheses. Biopterin was kindly donated by Drs. H. Hasegawa, T. Fukushima and M. Akino, Tokyo Metropolitan University and reduced to the tetrahydro form by a modification[1] of the method of Bobst and Viscontini (8). Sheep liver dihydropteridine reductase was purified according to the method of Kaufman and Levenberg (9) up to the step of the calcium phosphate gel treatment.

The activation and the assay of tryptophan 5-monooxygenase were performed as described previously (6), except that the activation was carried out at 30° for 30 min unless otherwise stated. Tyrosine and phenylalanine monooxygenase activities were determined under the same conditions as those for the assay of tryptophan 5-monooxygenase, except that L-tryptophan-1-^{14}C was replaced by 200 μM L-tyrosine-1-^{14}C or 200 μM L-phenylalanine-4-T, respectively, and the tyrosine monooxygenase reaction was carried out for 15 min. Tritium released from L-phenylalanine-4-T as THO or migrated to the adjacent position of the benzene ring was determined by a slight modification of the method of Guroff and Abramdwitz (10). L-Dopa-^{14}C formed from L-tyrosine-1-^{14}C was isolated by an alumina column according to the method of Nagatsu et al. (11) and counted. When the enzyme preparation contained L-tryptophan, the enzyme was dialyzed against 5 mM K-phosphate (pH 7.0) containing 5% glycerol, 3% ethanol and 20 mM $(NH_4)_2SO_4$ before the assay. The sucrose density gradient ultracentrifugation and the gel filtration on calibrated columns of Sephadex G-75 or G-100 were carried out according to the method of Martin and Ames (12) and of Andrews (13) or Siegel and Monty (14), respectively. Protein was determined by the method of Lowry et al. (15) with BSA as standard.

---

1. The authers are grateful to Drs. H. Hasegawa, T. Fukushima and M. Akino, Tokyo Metropolitan University, for making available their unpublished modified method for the reduction of biopterin.

RESULTS AND DISCUSSION

Purification --- All procedures for the purification of tryptophan 5-monooxygenase were carried out at 4°. The extraction of the enzyme from bovine pineal glands, the ammonium sulfate fractionation and the chromatography on a hydroxylapatite column were carried out as already described (6). All buffers used for the subsequent purification procedures contained 5% glycerol, 3% ethanol, 20 mM $(NH_4)_2SO_4$ and 0.1 mM L-tryptophan unless otherwise stated.

The concentrated hydroxylapatite fraction (6) obtained from 400 g of bovine pineal glands was applied on a column of P-cellulose (3.2 x 24 cm) which had been equilibrated with 5 mM K-phosphate (pH 7.0) and the enzyme was eluted by a linear gradient between 1.2 l each of 5 mM and 300 mM K-phosphate (pH 7.0). The active fractions (1-1.4 l) were combined, concentrated to 50 ml by ultrafiltration and dialyzed against 5 mM K-phosphate (pH 7.3)[2]. The dialyzed enzyme was then applied on a DEAE-cellulose column (2 x 36 cm) which had been equilibrated with 5 mM K-phosphate (pH 7.3). After the column was washed with 150 ml of the same buffer, a linear gradient elution was carried out between 500 ml each of 5 mM and 50 mM K-phosphate (pH 7.3). Fractions (12 ml each) were collected at a flow rate of 60 ml per hour. The enzyme was eluted into 2 peaks; a small peak in tubes No.9-20 was followed by a main activity peak in tubes No.30-55. The second active fractions were combined and made 80% saturated with respect to $(NH_4)_2SO_4$. The precipitate was collected by centrifugation and dissolved in a small amount of 20 mM K-phosphate (pH 7.0). The enzyme solution was then applied on a column of Sephadex G-200 (2 x 81 cm) which had been equilibrated with 20 mM K-phosphate (pH 7.0). Elution was carried out to the ascending direction at a flow rate of 10 ml per hour and 2 ml-fractions were collected. Active fractions (tubes No.76-88) were combined and concentrated to about 5 ml with the aid of a collodion bag. A summary of the purification is given in Table 1. The overall purification achieved was about 1,000-fold with a yield of 3%.

Properties of the enzyme --- Throughout the purification procedure phenylalanine monooxygenase activity remained associated with tryptophan 5-monooxygenase. Furthermore, when the purified enzyme was subjected to Sephadex G-100 column chromatography, both enzyme activities were eluted in the same single symmetrical peak and upon sucrose density gradient ultracentrifugation they moved down to an identical position. As shown in Fig. 1, both enzyme activities increased approximately 100-fold on the anaerobic preincubation in the presence of DTT and the time course of the activation was in parallel with each other both in the presence and absence of Fe^{2+}. When the enzyme was subjected to the

2. 5 mM K-phosphate (pH 7.3) used in the DEAE-cellulose column chromatography contained no $(NH_4)_2SO_4$.

Table 1. Purification of tryptophan 5-monooxygenase.

Enzyme preparations	Volume	Enzyme activity	Yield	Protein	Specific activity
	(ml)	(units)*	(%)	(mg)	(units/mg)
Homogenate	1,300	11,040	100	50,700	0.22
Digitonin Sup.	1,820	9,397	85	24,600	0.38
Amm. sulfate (25-57%)	270	7,171	65	5,990	1.20
Hydroxylapatite	125	3,548	32	1,100	3.23
P-Cellulose	46.5	2,521	23	393	6.42
DEAE-Cellulose	5.3	748.5	6.8	33.8	22.14
Sephadex G-200	4.6	400.2	3.6	1.7	235.40

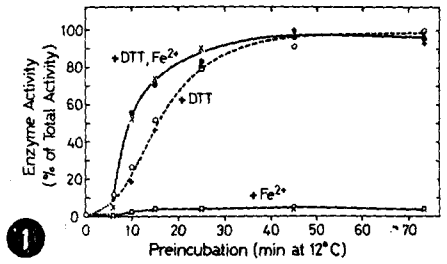
* One unit of enzyme activity was defined as the amount which catalyzed the hydroxylation of 1 nmole of L-tryptophan per min under the standard assay conditions.

heat treatment at various temperatures (45-65°) for 2 min or at 53° for various periods of time (1-10 min), both monooxygenase activities decreased to nearly the same extent. In good agreement with Lovenberg et al.(3), L-phenylalanine inhibited tryptophan 5-monooxygenase activity competitively with respect to L-tryptophan and vice versa. All experimental evidence described above thus suggested that hydroxylations of both L-tryptophan and L-phenylalanine were catalyzed by the single enzyme. On the other hand, the hydroxylation of L-tyrosine by a purified enzyme preparation[3] occurred at less than one-hundredth the velocity of the hydroxylation of L-tryptophan. Furthermore, L-tyrosine inhibited tryptophan and phenylalanine monooxygenase activities appreciably only at a high concentration with an approximate Ki value of 500 to 600 μM, suggesting that L-tyrosine is a much poorer substrate of this enzyme, if at all[4].

The apparent Km values of the enzyme for L-tryptophan and for L-phenyl-

3. The enzyme preparation was purified with the following modifications of the procedure: 1. The DEAE-cellulose step was omitted, 2. P-Cellulose and Sephadex G-200 were replaced by CM-Sephadex C-50 and Sephadex G-100, respectively. The specific activity of this preparation was 140 and 310 nmoles of tryptophan and phenylalanine hydroxylated, respectively, per min per mg of protein.

4. The trace activity of tyrosine hydroxylase in the purified preparation was first detected by Dr. S. Nagatsu, Aichi Gakuin University. The authers are grateful for his help. The possibility has not been excluded that the tyrosine hydroxylase activity was due to the contamination of other enzyme.



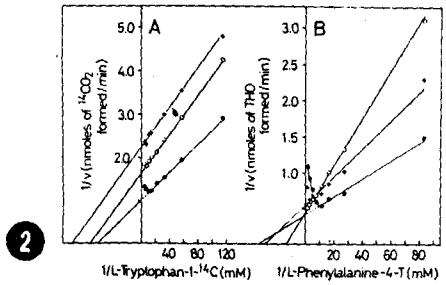


Fig. 1. Effect of preincubation on tryptophan and phenylalanine monooxygenase activities. The preincubation of the enzyme was carried out at 12° in the presence of either Fe^{2+}, DTT or Fe^{2+} plus DTT as described previously (6). At various time intervals an aliquot (50 µl) of the preincubated enzyme was removed and assayed for tryptophan (• and o) and phenylalanine (x and +) monooxygenase activities as described under "Methods".

Fig. 2. Tryptophan (A) and phenylalanine (B) monooxygenase activities as a function of substrate concentration. Enzyme activities were determined as described under "Methods" with the following modifications; 1. Reactions were carried out for 1 min instead of 3 min, 2. The concentration of DTT in the reaction mixture was reduced to 1.67 mM and an excess amount of dihydropteridine reductase and 0.1 mM NADPH were included, 3. Various concentrations of L-tryptophan-1-^{14}C (A) or L-phenylalanine-4-T (B) were used as substrate, 4. $DMPH_4$ (o, 0.89 mM), $6MPH_4$ (+, 0.36 mM) or BpH_4 (•, 0.77 mM) was used as a cofactor. The apparent Km values for L-tryptophan and L-phenylalanine estimated from this experiment were 14 and 70 µM (cofactor: $DMPH_4$), 11 and 37 µM (cofactor: $6MPH_4$) and 16 and 32 µM (cofactor: BpH_4), respectively.

---

alanine were approximately 16 and 32 µM, respectively, when 0.77 mM BpH_4 was used as a cofactor (Fig. 2). The Km for tryptophan did not change significantly with the structure of the pterin cofactor used but the value for phenylalanine increased to approximately 70 µM when BpH_4 was replaced by $DMPH_4$. The Vmax of phenylalanine hydroxylation was 2- to 3-fold higher than that of tryptophan hydroxylation. The apparent molecular weight and the Stokes radius of the enzyme were estimated to be approximately 30,000 and 25 A, respectively, by gel filtration on calibrated columns of Sephadex G-75 and G-100 and by ultracentrifugation in sucrose density gradients (Fig. 3).

Rat liver phenylalanine hydroxylase (E.C.1.14.3.1) has been shown to catalyze tryptophan hydroxylation as well (16) and the tryptophan 5-monooxygenase activity of that enzyme was shown to be stimulated by lysolecithin (17). However, the Km value for L-tryptophan and the Vmax of the tryptophan hydroxylation of rat liver phenylalanine hydroxylase were shown to be 1 mM and 0.1% of that of phenylalanine hydroxylation, respectively, even in the presence of lysolecithin (17). The molecular weights for the two major forms of phenylalanine hydroxylase have been reported to be 110,000 and 210,000 (18). In a preliminary experiment, phenylalanine hydroxylase was purified from the rat liver according to the method of Kaufman and Fisher (18) up to the step of second am-

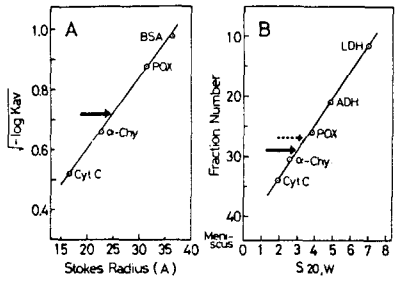


Fig. 3. Gel filtration on a Sephadex G-75 column (A) and sucrose density gradient ultracentrifugation (B) of tryptophan 5-monooxygenase. Analytical gel chromatography was carried out on a calibrated column (3.5 x 85 cm) of Sephadex G-75 equilibrated with 5 mM K-phosphate (pH 7.0) containing 5% glycerol, 3% ethanol, 20 mM $(NH_4)_2SO_4$ and 0.1 mM L-tryptophan. Six ml-fractions were collected at 4° at a flow rate of 12 ml per hour. Proteins with the following molecular weights and Stokes radii were employed as standards: BSA (Fraction V), 68,000, 36.4 A; horse radish peroxidase (POX), 44,000, 31.4 A; beef pancreatic α-chymotrypsinogen (α-Chy), 23,500, 22.6 A; and horse heart cytochrome C (Cyt C), 13,400, 16.5 A. The ultracentrifugation was carried out at 4° and 40,000 rpm for 42 hours in a sucrose gradient (5 to 16%) containing 50 mM K-phosphate (pH 7.0), 5% glycerol, 3% ethanol, 20 mM $(NH_4)_2SO_4$ and 0.1 mM L-tryptophan. After centrifugation 44 fractions (120 µl each) were collected from the bottom of the tube. Standard proteins used were rabbit skeletal muscle lactate dehydrogenase (LDH, S20,w = 7.0-7.3), horse liver alcohol dehydrogenase (ADH, S20,w = 4.82), horse radish peroxidase (POX, S20,w = 3.85), beef pancreatic α-chymotrypsinogen (α-Chy, S20,w = 2.58) and horse heart cytochrome C (Cyt C, S20,w = 1.9-2.1). The arrows indicate the peak of tryptophan 5-monooxygenase activity. The dotted arrow represents the presence of s smaller peak or a shoulder of the enzyme activity.

monium sulfate fractionation and this enzyme was shown to be activated only less than 2-fold on the anaerobic preincubation in the presence of DTT and Fe^{2+}. 6-Fluorotryptophan at 1 mM was found, in the present study, to produce greater than 90% inhibition of bovine pineal tryptophan 5-monooxygenase whereas the inhibitory effect on rat liver phenylalanine hydroxylase has been reported to be no more than 5% (19). The experimental evidence thus indicated that bovine pineal tryptophan 5-monooxygenase and rat liver phenylalanine hydroxylase are distinguishable by a number of criteria. The high activity of the pineal enzyme for tryptophan may suggest that the hydroxylation of tryptophan in the pineal gland is an important function of this enzyme.

REFERENCES

1. Grahame-Smith, D. G. (1967) Biochem. J. 105, 351-360.
2. Ichiyama, A., Nakamura, S., Nishizuka, Y., and Hayaishi, O. (1968) Advances in Pharmacology 6A, 5-17.
3. Lovenberg, W., Jequier, E., and Sjoerdsma, A. (1968) Advances in Pharmacology 6A, 21-36.
4. Ichiyama, A., Nakamura, S., Nishizuka, Y., and Hayaishi, O. (1970) J. Biol. Chem. 245, 1699-1709.
5. Friedman, P. P., Kappelman, A. H., and Kaufman, S. (1972) J. Biol. Chem. 247, 4165-4173.

6. Ichiyama, A., Hori, S., Mashimo, Y., Nukiwa, T., and Makuuchi, H. (1974) FEBS Letters 40, 88-91.
7. Harvey, J. A., and Gal, E. M. (1974) Science 183, 869-871.
8. Bobst, von A., and Viscontini, M. (1966) Helv. Chim. Acta 49, 875-884.
9. Kaufman, S., and Levenberg, B. (1959) J. Biol. Chem. 234, 2683-2688.
10. Guroff, G., and Abramdwitz, A. (1967) Anal. Biochem. 19, 548-555.
11. Nagatsu, T., Levitt, M., and Udenfriend, S. (1964) J. Biol. Chem. 239, 2910-2917.
12. Martin, R. G., and Ames, B. N. (1961) J. Biol. Chem. 236, 1372-1379.
13. Andrews, P. (1965) Biochem. J. 96, 595-606.
14. Siegel, L. M., and Monty, K. L. (1966) Biochim. Biophys. Acta 112, 346-362.
15. Lowry, O. H., Rosenbrough, N. J., Farr, A. L., and Randall, R. J. (1951) J. Biol. Chem. 193, 265-275.
16. Renson, J., Weissbach, H., and Udenfriend, S. (1962) J. Biol. Chem. 237, 2261-2264.
17. Fisher, D. B., and Kaufman, S. (1973) J. Biol. Chem. 248, 4345-4353.
18. Kaufman, S., and Fisher, D. B. (1970) J. Biol. Chem. 245, 4745-4750.
19. Kaufman, S., and Fisher, D. B. (1974) Molecular Mechanism of Oxygen Activation (Ed., O. Hayaishi), pp. 285-369, Academic Press, New York.

Reprinted from

Proc. Natl. Acad. Sci. USA
Vol. 75, No. 5, pp. 2140–2144, May 1978
Biochemistry

Organization of immunoglobulin heavy chain genes and allelic deletion model

(complementary DNA hybridization/myeloma/recombination of variable and constant region genes)

TASUKU HONJO AND TOHRU KATAOKA

Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo 113, Japan

Communicated by Osamu Hayaishi, February 7, 1978

ABSTRACT **We have assessed the number of times the gene sequence encoding constant regions of mouse immunoglobulin heavy chains γ1, γ2a, and γ3 are represented in the mouse genome by hybridization kinetic analysis. All three genes are present at one copy each per haploid genome in normal tissues and myelomas producing IgM or IgG3. IgG1-producing myelomas, however, contain 1 copy each of the γ1 and γ2a genes and 0.5 copy of the γ3 gene per haploid genome. IgG2b-producing myelomas contain 1 copy of the γ2a gene and 0.5 copy each of the γ1 and γ3 genes per haploid genome. IgG2a-producing myelomas contain 1 copy of the γ2a gene and 0.5 copy each of the γ1 and γ3 genes per haploid genome. In myelomas producing IgA, all three γ genes are represented 0.5 time per haploid genome. In order to account for the results we propose an allelic deletion model: (*i*) The specific deletion of heavy chain constant region genes accompanies the recombination of a variable region gene to a constant region gene. (*ii*) The portion of the chromosome that resides between two joining sequences is excised out of the chromosome. (*iii*) The recombination occurs on one of the alleles. Based on this model we also propose that heavy chain genes are arranged on one chromosome in the following order; variable region genes, unknown spacer sequence, μ, γ3, γ1, γ2b, γ2a, and α.**

The genetic control of immunoglobulin synthesis has been a puzzle to molecular biologists as well as to immunologists. The organization of immunoglobulin genes is a key to those unresolved problems unique to immunology, such as the two gene–one polypeptide theory, allelic exclusion, and the origin of antibody diversity.

There are at least seven classes and subclasses of immunoglobulins in the mouse—IgM, IgD, IgG1, IgG2a, IgG2b, IgG3, and IgA—which differ from each other in the amino acid sequences of the constant regions (C) of the heavy (H) chains. Genetic studies have shown that the H chain C genes of the mouse are clustered to such a narrow segment of one chromosome that there have been no recombinants found among 2069 progenies derived from two crosses (1, 2). These C_H genes share a family of variable region (V) genes (3–5) that resides on the same chromosome and is separated from the C_H genes by unknown spacer sequences (6).

Dreyer and Bennett (7) have proposed that V and C are encoded by two separate genes which eventually join to form a single gene. Extensive kinetic studies using probes derived from purified light (L) chain mRNA have produced general agreement in favor of the V–C recombination hypothesis (8–16).

In contrast, there have been few reports of molecular studies on organization of H chain genes (17, 18). This is mainly because purification of H chain mRNA is much more difficult than that of L chain mRNA. Recently we have succeeded in purifying H chain mRNA from mouse myelomas of several subclasses by immunoprecipitation of polysomes (19). Using these probes we have measured the number of H chain gene sequences and shown deletion of H chain c genes in mouse myelomas that produce immunoglobulins of specific classes. Based on these results we will propose an allelic deletion model for the mechanism of V_H–C_H recombination. The H chain c genes are proposed to be aligned on one chromosome in the order V_H, spacer, μ, γ3, γ1, γ2b, γ2a, and α.

MATERIALS AND METHODS

Myeloma tumor lines used in the present study were kindly supplied by M. Potter (National Institutes of Health), T. Kishimoto (Osaka University), M. Kawakami (Kitazato University), and S. Migita (Kanazawa University). Myeloma tumors were propagated and harvested as described (9).

H chain mRNAs were purified from MOPC 31C (IgG1), J606 (IgG3), MPC 11 (IgG2b), and HOPC 1 (IgG2a) myelomas by M. Ono as described (19). The purities of MOPC 31C, HOPC 1, MPC 11, and J606 mRNAs are 97%, 54%, 92%, and 99%, respectively, as determined by hybridization kinetic analysis (9). More detailed characterization will be described elsewhere. Mouse globin and κ L chain mRNAs were prepared from mouse reticulocytes and MOPC 41 myeloma, respectively, as described (9). [^{3}H]cDNAs complementary to mRNAs encoding globin, κ L chain, and H chain were synthesized and purified as described (9, 20). H chain cDNAs derived from MOPC 31C and J606 mRNA are about 1500 bases long; H chain cDNA from HOPC 1 mRNA is about 1800 bases long. The specific radioactivities of globin, κ L chain, and H chain [^{3}H]cDNAs thus prepared were 10, 10, and 16 cpm/pg, respectively. DNAs from tumors and normal tissues were prepared as described (21). DNA was digested twice with RNase (100 μg/ml, 2 hr at 37°) and treated with alkali (0.2 M, 60 min at 37°) before use. DNA was sheared to 800 base-pairs long by sonication.

Hybridization of radioactive probes with cellular DNA was performed as described (9, 14) except that ^{14}C-labeled DNA derived from MOPC 31C myeloma cells was added to all the hybridization reaction mixtures to 610 cpm/mg. The concentrations of cellular DNA, [^{3}H]cDNA, and Na^+ were 9.4 mg/ml, 350 pg/ml, and 1 M, respectively. The hybrids formed were assayed by S1 nuclease digestion.

RESULTS AND DISCUSSION

Characterization of Probes. H chain mRNAs were purified from MOPC 31C (γ1), HOPC 1 (γ2a), MPC 11 (γ2b), and J606

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 and C, variable and constant regions; H and L chains, heavy and light chains; cDNA, synthetic DNA complementary to mRNA; C_0t value, product of concentration (mol/liter) of nucleotide sequences of DNA and time (sec) of incubation.

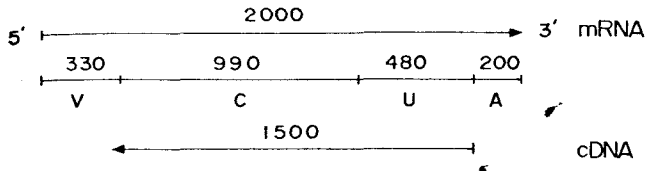


FIG. 1. Diagram of H chain mRNA and its cDNA. V, variable region sequence; C, constant region sequence; U, untranslated sequence; A, poly(A). Numbers are numbers of bases.

(γ3) myelomas by immunoprecipitation of polysomes (19). The mRNA preparations obtained ran as homogeneous bands upon polyacrylamide gel electrophoresis in formamide, with a molecular weight of 700,000 (i.e., 2000 bases).

The cDNA preparations used for the present study were about 1500 bases long as determined by polyacrylamide gel electrophoresis in formamide. Because we did not see any difference in the extent of hybridization of cDNA between homologous and heterologous H chain mRNAs of the identical subclass, the cDNA preparations represent most of the C gene sequences but not of the V gene sequences. A diagram of H chain mRNA and its cDNA is shown in Fig. 1. Nucleotide sequence analyses of MOPC 31C cDNA inserted in plasmids indicated that the untranslated sequence of about 500 nucleotides is present at the 3' end of mRNA (unpublished data).

Reiteration Frequency of the γ1 Gene. MOPC 31C [^{3}H]-cDNA was hybridized to a vast excess of cellular DNA derived from MOPC 31C myeloma. Reassociation of mouse unique DNA was followed by the addition of a trace amount of ^{14}C-labeled MOPC 31C DNA. All C_0t analyses were carried out in similar double-label experiments under the identical conditions.

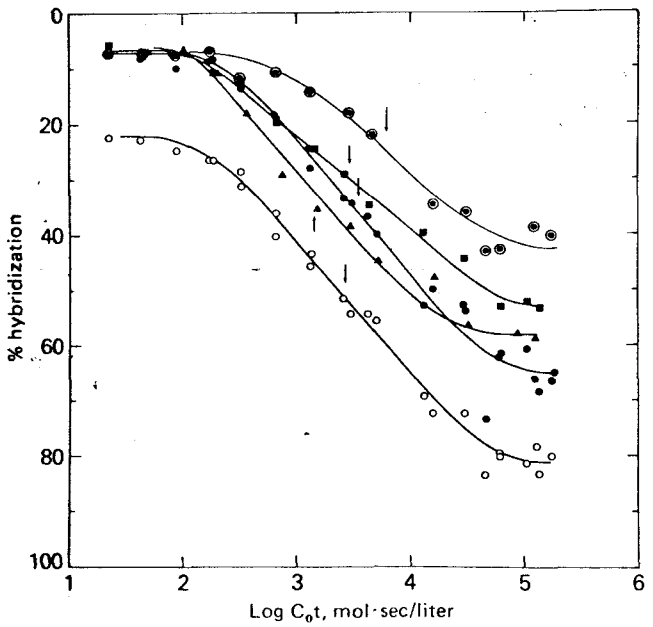


FIG. 2. Hybridization kinetic analyses of MOPC 31C cDNA with total cellular DNA from MOPC 31C (IgG1) and MOPC 511 (IgA) myelomas. MOPC 31C [^{3}H]cDNA was hybridized to MOPC 31C DNA or MOPC 511 DNA. In separate experiments, κ L chain [^{3}H]cDNA and globin [^{3}H]cDNA were hybridized to total cellular DNA from MOPC 31C myeloma; these reactions were carried out as above except that concentrations of globin [^{3}H]cDNA and κ chain [^{3}H]cDNA were 540 and 800 pg/ml, respectively. Hybridization profiles of MOPC 31C [^{14}C]DNA in three experiments were superimposable. Arrows indicate $C_0t_{1/2}$ points. ○, MOPC 31C [^{14}C]DNA; ●, MOPC 31C cDNA with MOPC 31C DNA; ◉, MOPC 31C cDNA with MOPC 511 DNA; ■, κ chain cDNA; ▲, globin cDNA.

Table 1. Reiteration frequency of the γ1 gene

Origin of DNA				
	H chain class	Copy/haploid		
Myeloma	produced	γ1	Lκ	Globin
MOPC 104E	μ	1.0		
J606	γ3	1.1		
MOPC 31C	γ1	0.79	0.93	1.9
MOPC 70A	γ1	0.82		
MC 101	γ1	1.03		
MPC 11	γ2b	0.42	1.1	
NP 2	γ2b	0.39	1.1	
MOPC 141	γ2b	0.59		
UPC 10	γ2a	0.60		
RPC 5	γ2a	0.56		
MOPC 511	α	0.48		
MOPC 315	α	0.52		
MOPC 41	None	0.53	1.0	1.8
RPC 20	None	0.44		
Liver	—	1.3	1.0	1.9
Spleen	—	1.0		
Kidney	—	0.96		
Thymus	—	1.3		
Newborn mice	—	1.1	1.0	

As shown in Fig. 2, γ1 cDNA hybridized with a $C_0t_{1/2}$ value of 3500, whereas that of mouse unique sequences is 2700. Comparison of these $C_0t_{1/2}$ values yields a reiteration frequency of 0.8 copy per haploid genome for the γ1 gene sequence in MOPC 31C DNA (22). We have tested the congruity of the cDNA·DNA hybrids by thermal stability to S1 nuclease digestion (14). The γ1 cDNA hybrid formed with MOPC 31C DNA showed a sharp melting profile with a t_m value of 90.6°, indicating that congruently matched duplexes were formed. As control experiments, κ L chain cDNA derived from MOPC 41 mRNA and globin cDNA were hybridized to the same batch of MOPC 31C DNA. $C_0t_{1/2}$ values for the κ chain cDNA and globin cDNA were 2900 and 1400, respectively, indicating that the κ chain and globin genes are represented 0.93 and 1.9 times per haploid genome, respectively, in agreement with previous reports (8–14, 21). Similar experiments were carried out with DNA derived from many myelomas and normal tissues. As summarized in Table 1, the γ1 gene sequence is present as 1 copy per haploid genome in IgG1-, IgM-, and IgG3-producing myelomas as well as in normal mouse tissues.

On the other hand, when the γ1 cDNA was hybridized to DNA derived from an IgA-producing myeloma (MOPC 511), the hybridization kinetics were quite different from those obtained for MOPC 31C DNA. As shown in Fig. 2, a $C_0t_{1/2}$ value of 6200 was obtained, compared to 3000 for mouse unique copy DNA (not shown). The results yield a reiteration frequency of 0.48 copy per haploid genome. The extent of hybridization reached approximately 42% which is smaller than the value (65%) obtained with MOPC 31C DNA. The possibility that reduction in the extent of hybridization is due to a partial deletion of γ1 sequence was excluded by the following experiments. First, cDNA preparations of different length (500, 1000, and 1500 bases) gave the identical extent of hybridization. Second, the extent of hybridization reached about 65% when the ratio of DNA to cDNA was doubled.

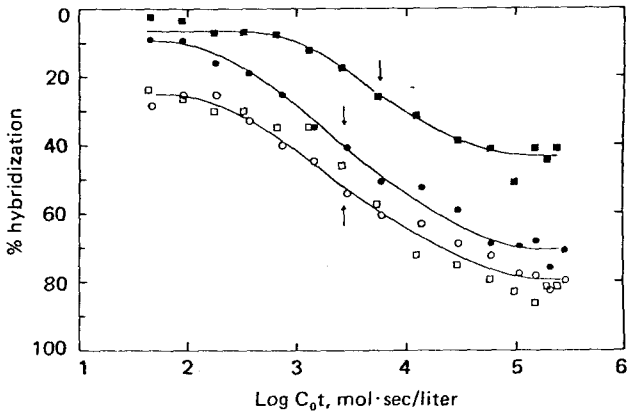


FIG. 3. Hybridization kinetic analyses of J606 cDNA with total cellular DNA from J606 and MOPC 70A myelomas. Arrows indicate $C_0t_{1/2}$ points. ○, J606 [^{14}C]DNA; ●, J606 [^{3}H]cDNA with J606 DNA; □, MOPC 70A [^{14}C]DNA; ■, J606 [^{3}H]cDNA with MOPC 70A DNA.

In nine lines of myeloma producing IgG2a, IgG2b, IgA, or Bence Jones protein, the number of $\gamma 1$ gene copy ranged from 0.39 to 0.60 (Table 1). The results indicate that the average number of the $\gamma 1$ gene copy is reduced to 0.5 copy per haploid genome in these myelomas. Our technique is sensitive enough to distinguish 1 copy per haploid from 2 copies per haploid because the number of gene copies encoding globin and κ chain remained constant regardless of the reduction in the $\gamma 1$ gene copy. Obviously, the results can be interpreted in two ways. One possibility is that chromosomal abnormalities in these myelomas happen to have decreased the apparent reiteration frequency of the $\gamma 1$ gene (23). The alternative possibility, which we prefer, is that the $\gamma 1$ gene is specifically deleted from one of the alleles in mouse myelomas that produce immunoglobulins of the defined classes (i.e., IgG2a, IgG2b, and IgA).

Reiteration Frequency of the $\gamma 3$ Gene. In order to test the possibilities mentioned above, we assessed the number of times the gene sequences corresponding to the $\gamma 3$ cDNA probe are represented in various myelomas by hybridization kinetic analysis. Typical examples of such experiments are shown in Fig. 3, in which J606 [^{3}H]cDNA was hybridized to a vast excess of cellular DNA derived from J606 (IgG3) and MOPC 70A (IgG1) myelomas, respectively. The results are quite clear. The $C_0t_{1/2}$ values for DNAs from J606 and MOPC 70A are 2700 and 5800, respectively, compared to 2700 for unique copy mouse DNA. These values are equivalent to reiteration frequencies of 1.0 and 0.46 per haploid genome for DNAs from J606 and MOPC 70A, respectively.

Similar experiments were carried out with DNAs derived from 11 other myelomas and 3 tissues (Table 2). IgM- and IgG3-producing myelomas, as well as normal tissues, contain 1 copy of the $\gamma 3$ gene sequence per haploid genome whereas IgG1-, IgG2a-, IgG2b-, and IgA-producing myelomas contain only 0.5 copy per haploid genome. The maximal extents of hybridization decreased in concordance with reduction in the reiteration frequency. It is worth emphasizing that IgG1-producing myelomas MOPC 31C, MOPC 70A, and MC 101 have lost 0.5 copy of the $\gamma 3$ gene sequence per haploid genome. Nevertheless, they retain 1 copy of the $\gamma 1$ gene sequence per haploid. Because both $\gamma 1$ and $\gamma 3$ gene sequences have been shown to be located on a small segment of one chromosome, it is unlikely that the reduction in the number of the $\gamma 3$ gene copy is attributable to general chromosomal abnormalities in these myelomas.

Table 2. Reiteration frequency of the $\gamma 3$ and $\gamma 2a$ genes

Origin of DNA			
Myeloma	H chain class produced	Copy/haploid $\gamma 3$	Copy/haploid $\gamma 2a$
MOPC 104E	μ	0.85	0.94
J606	$\gamma 3$	1.0	1.20
FlOPC 21	$\gamma 3$	0.82	0.93
MOPC 31C	$\gamma 1$	0.67	1.13
MOPC 70A	$\gamma 1$	0.52	—
MC 101	$\gamma 1$	0.54	—
MPC 11	$\gamma 2b$	0.52	1.36
NP 2	$\gamma 2b$	0.5	—
MOPC 141	$\gamma 2b$	0.63	—
UPC 10	$\gamma 2a$	0.52	1.2
RPC 5	$\gamma 2a$	0.68	—
HOPC 1	$\gamma 2a$	—	1.2
MOPC 511	α	0.46	0.56
MOPC 315	α	0.37	0.60
Liver	—	1.18	—
Spleen	—	—	1.0
Kidney	—	0.88	1.0
Newborn mice	—	0.9	—

Reiteration Frequency of the $\gamma 2a$ and $\gamma 2b$ Genes. Determination of the copy number of the $\gamma 2$ gene represented in the myeloma genome provided further evidence for the specific deletion of the γ genes in mouse myelomas (Table 2). Approximately 1 copy of the $\gamma 2a$ gene sequence is present per haploid genome in DNAs derived from normal tissues and myelomas producing IgM, IgG3, IgG1, IgG2b, and IgG2a. Only 0.5 copy of the $\gamma 2a$ gene is present per haploid genome in DNA derived from myelomas producing IgA.

Preliminary results indicate that the $\gamma 2b$ gene is represented once per haploid genome in DNAs derived from spleen and myelomas producing IgM, IgG3, IgG1, and IgG2b whereas it is reduced to approximately 0.5 copy per haploid genome in DNAs from myelomas producing IgA.

Summary of γ Gene Deletions. Taking all these results together, we are inclined to conclude that the reduction in the number of the γ gene copy in mouse myelomas is due to the specific deletion of the particular genes from one of the alleles. A summary of our interpretation is shown in Table 3. Assuming that the H chain genes are arranged on a chromosome in the order V_H genes, spacer sequence, μ gene, $\gamma 3$ gene, $\gamma 1$ gene, $\gamma 2b$ gene, $\gamma 2a$ gene, and α gene, a whole set of the results can be easily explained by deletion of a chromosomal segment from one of the alleles. Apparently, the deletion starts next to the C_H gene expressed in each myeloma, suggesting that the deletion

Table 3. Summary of deletion of γ genes

Origin of DNA	Copy/haploid $\gamma 3$	$\gamma 1$	$\gamma 2b$	$\gamma 2a$
IgM producer	1	1	1	1
IgG3 producer	1	1	1	1
IgG1 producer	0.5	1	1	1
IgG2b producer	0.5	0.5	1	1
IgG2a producer	0.5	0.5	0.5	1
IgA producer	0.5	0.5	0.5	0.5
Nonlymphatic tissues	1	1	1	1

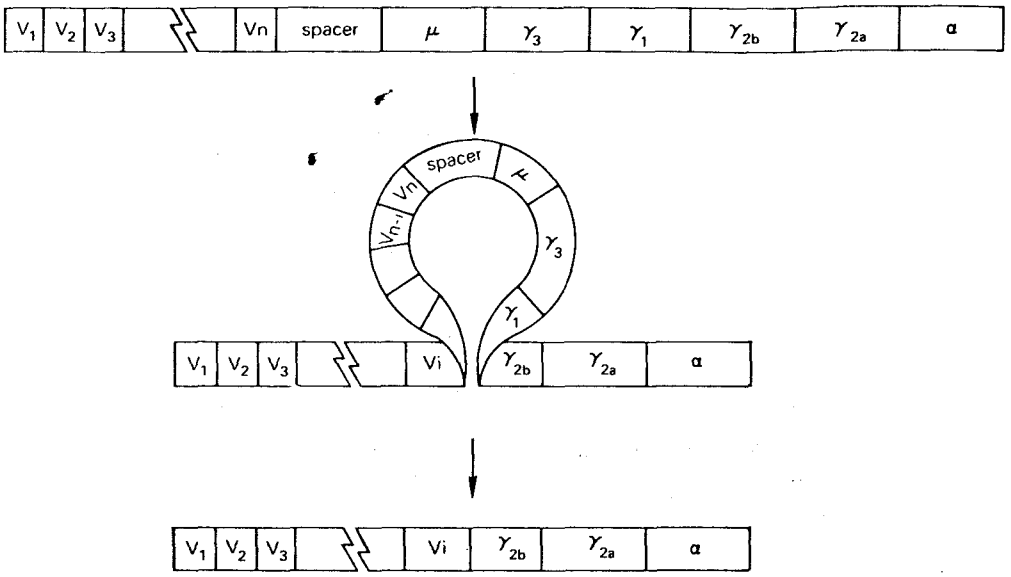


FIG. 4. Allelic deletion model for V_H–V_C recombination. The diagram shows recombination of a V_i sequence with the γ2b gene, resulting in deletion of the μ, γ3, and γ1 genes.

of the C_H genes is related to expression of the remaining C_H gene.

Allelic Deletion Model. In order to put all these results together, we propose a model that we call the "allelic deletion model." We assume that the specific deletion of C_H gene sequences accompanies the V–C gene recombination which takes place only on one of the alleles, as shown in Fig. 4. Upon joining of the V_H and C_H gene sequences the portion of the chromosome that resides between the two sequences is excised out of the chromosome. The deletion of the γ genes in myelomas producing the specific classes of immunoglobulin reflects the relative location of the class-specific C_H genes that have undergone recombination with a V_H gene. Given the deletion profile shown in Table 3, it is possible to assume that C_H genes are arranged in the order (μ, γ3), γ1, γ2b, γ2a, and α genes. It is also reasonable to put the μ gene between the V_H sequences and the γ3 sequence. Otherwise, the γ3 gene should be deleted in the IgM-producing myeloma.

We think that a given V_H gene recombines successively with different C_H genes during immunodifferentiation of a single lymphocyte. Inasmuch as C_H gene deletion accompanies the V–C recombination, differentiation proceeds irreversibly from an IgM-producer to an IgG- or IgA- producer according to a linear arrangement of C_H genes. We have no information concerning the location of the δ gene because there has been no IgD-producing mouse myeloma available. We speculate that the δ gene is located between the μ gene and the γ3 gene because the δ gene is expressed prior to the γ genes but after the μ gene during the course of differentiation (24, 25). But we reserve the possibility that the δ gene is an exception to the model.

The proposed order of C_H genes is compatible with recent genetic studies using rabbit allotypes (26–28), which indicate that the μ gene is closest to the V_H genes and that the α gene is farthest from the V_H genes. Preud'homme *et al.* (29) reported that MPC 11 myeloma cells (IgG2b producer), when treated with mutagens, gave rise to mutant clones that produced γ2a H chains. This report is also in agreement with the proposed order of the γ2b and γ2a genes. Lieberman and Potter (30) proposed the order γ2a, γ2b, γ1, and α genes on the basis of the assumption that an H chain gene allele of the Japanese wild mouse has arisen from recombination between two alleles of inbred strains BALB/c and C57BL/6. However, different conclusions are drawn if different combinations are chosen as parent alleles.

Concordance of the Allelic Deletion Model with Immunological Phenomena. Any model for the immunoglobulin genes has to explain such unique immunological observations as allelic exclusion and switch of the immunoglobulin class synthesized in a single lymphocyte. Genetic studies using allotype markers have shown that only one of the alleles of immunoglobulin genes is expressed in activated B lymphocytes (31–33). Genetic information for the V_H and C_H which are genetically linked on one parental chromosome are coordinately expressed, which is referred to as *cis* expression (34–36). Apparently, the allelic deletion model postulates allelic exclusion and *cis* expression by nature. Tonegawa *et al.* (37) reported that rearrangement of κ chain genes occurs on both alleles. It remains to be seen whether different mechanisms operate on the L and H gene systems.

The switch within a clone from μ chain synthesis to γ chain synthesis is well established as the normal process by which B lymphocytes begin to secrete IgG, and this easily explains the simultaneous presence of both IgM and IgG on the cell surface (38–40). It is not clear, however, whether this switch may include a period in which a single cell simultaneously synthesizes μ and γ chains and, in particular, μ and γ mRNAs. According to the allelic deletion model the successive recombination of a V_H gene sequence, first with a μ gene and then with a γ gene, can account for the shift of the immunoglobulin class synthesized in a lymphocyte. The switch to the reverse direction has not been found and is very difficult to explain by this model. In view of the relatively long half-life of immunoglobulin proteins (96 hr) and mRNAs (14 hr) (41, 42), we do not have to postulate the concurrent transcription of μ and γ gene sequences within a single lymphocyte. A similar explanation may be applicable to the simultaneous presence of IgM and IgD on the cell surface (43).

Although no experimental data are available, all the possible mechanisms thus far proposed to account for the V–C recombination include (*i*) the copy-insertion model (5, 44), (*ii*) the translocation model (45), (*iii*) the inversion model (37), and (*iv*)

the looping-out excision model (46). None of the models except the looping-out excision model can explain our results.

We are grateful to Dr. Y. Mano for continuous encouragement, to Drs. M. Ono and M. Kawakami for preparation of purified H chain mRNAs, and to Drs. P. Leder, J. Gruber, and J. W. Beard for a generous gift of reverse transcriptase. This investigation was supported by grants from the Jane Coffin Child Memorial Fund for Cancer Research, the Matsunaga Science Foundation, and the Ministry of Education, Science, and Culture of Japan.

1. Lieberman, R. & Potter, M. (1966) *J. Mol. Biol.* **18**, 516–528.
2. Potter, M. & Lieberman, R. (1967) *Adv. Immunol.* **7**, 91–145.
3. Fudenberg, H. H., Wang, A. C., Pink, J. R. L. & Levin, A. S. (1971) *Ann. N.Y. Acad. Sci.* **190**, 501–506.
4. Wang, A. C., Gergely, J. & Fudenberg, H. H. (1973) *Biochemistry* **12**, 528–534.
5. Sledge, C., Fair, D. S., Black, B., Kruegar, R. G. & Hood, L. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 923–927.
6. Eichmann, K. (1975) *Immunogenetics* **2**, 491–506.
7. Dreyer, W. J. & Bennett, J. C. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 864–869.
8. Faust, C. H., Diggelman, H. & Mach, B. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2491–2495.
9. Honjo, T., Packman, S., Swan, D. & Leder, P. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3659–3663.
10. Stavnezer, J., Huang, R. C. C., Stavnezer, E. & Bishop, J. (1974) *J. Mol. Biol.* **88**, 43–63.
11. Tonegawa, S., Steinberg, C., Dube, S. & Bernardini, A. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4027–4031.
12. Rabbitts, T. H., Jarvis, J. M. & Milstein, C., (1975) *Cell* **6**, 5–12.
13. France, M.-G., Aellen, M.-F., Briand, P.-A., Faust, C. H., Vassalli, P. & Mach, B. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 727–731.
14. Honjo, T., Packman, S., Swan, D. & Leder, P. (1976) *Biochemistry* **15**, 2780–2785.
15. Leder, P., Honjo, T., Seidman, J. & Swan, D. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 855–862.
16. Tonegawa, S. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 203–207.
17. Premkumar; E., Shoyab, M. & Williamson, A. R. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 99–103.
18. Bernardini, A. & Tonegawa, S. (1974) *FEBS Lett.* **41**, 73–77.
19. Ono, M., Kondo, T., Kawakami, M. & Honjo, T. (1976) *J. Biochem. (Tokyo)* **81**, 949–954.
20. Monahan, J. J., Harris, S. E., Woo, S. L. C., Robberson, D. L. & O'Malley, B. W. (1976) *Biochemistry* **15**, 223–233.
21. Packman, S., Aviv, H., Ross, J. & Leder, P., (1972) *Biochem. Biophys. Res. Commun.* **49**, 813–819.
22. Britten, R. J. & Kohne, D. E. (1968) *Science* **161**, 529–540.
23. Yoshida, T. H., Imai, H. T. & Potter, M. (1968) *J. Natl. Cancer Inst.* **41**, 1083–1095.
24. Vitetta, E. S., Cambier, J., Forman, J., Kettman, J. R., Yuan, D. & Uhr, J. W. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 185–191.
25. Parkhouse, R. M. E., Abney, E. R., Bourgois, A. & Willcox, H. N. A. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 193–200.
26. Mage, R. G., Young-Cooper, G. O. & Alexander, C. (1971) *Nature New Biol.* **230**, 63–64.
27. Kindt, T. J. & Mandy, W. J. (1972) *J. Immunol.* **108**, 1110–1113.
28. Hamers-Casterman, C. & Hamers, R. (1975) *Immunogenetics* **2**, 597–603.
29. Preud'homme, J., Birshtein, B. K. & Scharff, M. D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1427–1430.
30. Lieberman, R. & Potter, M. (1969) *J. Exp. Med.* **130**, 519–541.
31. Pernis, B. G., Chiappino, G., Kelus, A. S. & Gell, B. G. H. (1965) *J. Exp. Med.* **122**, 853–876.
32. Cebra, J. J., Colberg, J. E. & Dray, S. (1966) *J. Exp. Med.* **123**, 547–558.
33. Mäkelä, O. & Cross, A. W. (1970) *Prog. Allergy* **14**, 145–207.
34. Landucci-Tosi, S., Mage, R. G. & Dubiski, S. (1970) *J. Immunol.* **104**, 641–647.
35. Kindt, T. J., Mandy, W. J. & Todd, C. W. (1970) *Biochemistry* **9**, 2028–2032.
36. Knight, K. L. & Hanly, W. C. (1975) in *Contemporary Topics in Molecular Immunology*, eds. Inman, F. P. & Mandy, W. J. (Plenum, New York), Vol. 4, p. 55–88.
37. Tonegawa, S., Hozumi, N., Matthyssens, G. & Schuller, R. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 877–889.
38. Kincade, P. W., Lawton, A. R., Bockman, D. E. & Cooper, M. D. (1970) *Proc. Natl. Acad. Sci. USA* **67**, 1918–1925.
39. Pierce, C. W., Solliday, S. M. & Asofsky, R. (1972) *J. Exp. Med.* **135**, 675–697.
40. Jones, P. P., Tacier-Eugster, H. & Herzenberg, L. A. (1974) *Ann. Immunol. (Poznan)* **125C**, 271–276.
41. Melchers, F., Lafleur, L. & Anderson, J. (1974) in *Control of Proliferation in Animal Cells*, eds. Clarkson, B. & Baserga, R. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) Vol. 1, pp. 393–410.
42. Cowan, N. J. & Milstein, C. (1974) *J. Mol. Biol.* **82**, 469–481.
43. Pernis, B., Forni, L. & Luzzati, A. L. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 175–183.
44. Dreyer, W. J., Gray, W. R. & Hood, L. (1968) *Cold Spring Harbor Symp. Quant. Biol.* **32**, 353–367.
45. Gally, J. A. & Edelman, G. M. (1970) *Nature* **227**, 341–342.
46. Kabat, D. (1972) *Science* **175**, 134–140.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 255, No. 11, Issue of June 10, pp. 5291-5295, 1980
Printed in U.S.A.

Untranslated Immunoglobulin κ Light Chain mRNA in a λ Light Chain-producing Mouse Myeloma, MOPC104E*

(Received for publication, November 8, 1979)

Tohru Kataoka, Masao Ono,‡ Masaya Kawakami,‡ Yoji Ikawa,§ Mariko Aida,§ Yoshitake Mano, and Tasuku Honjo¶

From the Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo, ‡the Molecular Biology Laboratory, School of Medicine, Kitasato University, Sagamihara, and §the Department of Viral Oncology, Cancer Institute, Tokyo, Japan

Fourteen clones were isolated in culture from a mouse myeloma, MOPC104E. All clones had κ and λ types of light chain mRNAs in approximately equimolar quantity as assayed by hybridization with specific complementary DNA (cDNA). However, the myeloma produces and secretes only λ-type light chain protein. Both κ- and λ-type mRNAs in these clones were indistinguishable from κ- and λ-type mRNAs of other myelomas with respect to (*a*) adsorption to oligo-(dT) cellulose, (*b*) molecular size (12.6 S), and (*c*) thermal stability of the hybrids formed with corresponding cDNA. The κ chain mRNA of MOPC104E cells, however, was translated very inefficiently both *in vivo* and *in vitro*, whereas the λ chain mRNA was translated efficiently. These results indicate that each cell of MOPC104E myeloma synthesizes a crippled κ chain mRNA in addition to a normal λ chain mRNA.

Several lines of evidence suggest that an immunoglobulin-secreting lymphocyte produces a single class of heavy chain and a single type of light chain at a time (1-3). Since a myeloma produces and secretes a homogeneous immunoglobulin (3), mRNA derived from the myeloma has been thought to encode the class and type of the immunoglobulin secreted. This point has been proven for many mouse myelomas by purification of L chain[1] mRNAs (4-9). We have shown, however, that a mouse myeloma MOPC104E, which produces a λ-type L chain as a constituent of immunoglobulin M (11), contains both κ and λ type L chain mRNAs in an equal quantity (10). We report here that several cloned cell lines of MOPC104E contain both L_κ and L_λ chain mRNAs in a similar quantity. The L_κ chain mRNA, however, seems to be translated very inefficiently both *in vivo* and *in vitro*.

EXPERIMENTAL PROCEDURES

Materials—Mouse myelomas MOPC41 and RPC20 were kindly donated by Dr. M. Potter, National Institutes of Health, and maintained by subcutaneous transplantation. Cultured MOPC104E cell line was donated by Dr. Y. Namba of Kyoto University. Purified reverse transcriptase from avian myeloblastosis virus was obtained from Dr. P. Leder of the National Institutes of Health. Rabbit antisera against L_κ or L_λ chain protein of mouse were raised and donated by Dr. S. Migita of Kanazawa University. IgG was further purified from the antisera by ammonium sulfate fractionation and DEAE-cellulose chromatography (12). Culture media and radiochemicals were purchased from Gibco and New England Nuclear Co., respectively. Other reagents unless specified were from the Wako Pure Chemicals Co., Osaka, Japan.

Cell Cloning—MOPC104E cells were first adapted to grow in HAM-F12 medium supplemented with 10% FCS and kanamycin, 60 μg/ml (Banyu Pharmaceutical Co., Tokyo). Subculture frequency was twice a week, with 1×10^5 cells/ml seeded. MOPC104E cells thus maintained were resuspended in fresh medium, and the cell suspension was diluted so that one drop of 0.2 ml contained 0.8 cells (by calculation). One drop was poured into each well of a microtiter plate (Falcon, Microplate II). Thirty-two stable clones were obtained. Fourteen clones were randomly selected for the present study.

Synthesis of Type-specific cDNAs—^{3}H-Labeled L_κ chain cDNA was synthesized by avian myeloblastosis virus reverse transcriptase and [^{3}H]dCTP (23 Ci/mmol) from purified MOPC41 L_κ chain mRNA as described (6). Long cDNA species with an average length of 900 nucleotides were selected by alkaline sucrose density gradient centrifugation (13). Two kinds of L_λ cDNA were used. cDNA synthesized from purified MOPC104E L chain mRNA was hybridized with purified L_λ chain mRNA of RPC20 myeloma (8), and the resultant hybrids were selected by hydroxylapatite chromatography (14). The hybridized cDNA was used as L_λ cDNA after alkaline treatment. L_λ cDNA synthesized against purified L_λ chain mRNA of RPC20 myeloma (8) was also used and gave identical results. The type specificity of the cDNAs was shown previously (10).

Preparation of Intra- and Extracellular Proteins—Cultured MOPC104E cells were suspended in 16.4 ml of L-15 medium depleted of leucine at 2×10^6 cells/ml and incubated at 37°C for 5.5 h in the presence of 25 μCi of [U-^{14}C]leucine (311 mCi/mmol). After centrifugation at 2000 rpm for 7 min, the supernatant was used to isolate extracellular proteins. The precipitated cells were washed twice with phosphate-buffered saline, suspended in 25 mM Tris-HCl, pH 7.4, 25 mM NaCl, 5 mM $MgCl_2$, 0.2 M sucrose, and 1% Triton X-100, homogenized with a Potter's homogenizer, and centrifuged at 100,000 × *g* at 4°C for 1 h. The resultant supernatant was used to isolate intracellular proteins.

Immunoprecipitation—Anti-L_κ chain IgG (22 mg) or anti-L_λ chain IgG (18 mg) were separately coupled to 10 ml of Sepharose 4B activated by cyanogen bromide in 0.2 M potassium phosphate, pH 6.4, at 4°C overnight (15, 16). The residual activated groups were blocked by 1 M ethanolamine/HCl, pH 8.7. Immunoprecipitation was performed in 10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, and 10 mM L-leucine. Labeled proteins and various amounts of anti-κ or anti-λ Sepharose suspensions were included in 1.0 ml of the reaction mixtures. After incubation at room temperature for 1 h, Sepharose particles were centrifuged through a solution of 1.2 M sucrose containing the reaction buffer to reduce the background (17). The centrifugation was repeated three times. The particles were trapped on a GF/C filter (Whatman), washed, and dried under a lamp. Proteins bound to the Sepharose were solubilized

* This investigation was supported in part by grants from the Ministry of Education, Science, and Culture, Japan, and from the Jane Coffin Childs Memorial Fund for Cancer Research. 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. Section 1734 solely to indicate this fact.

¶ To whom all correspondence should be addressed at the Department of Genetics, Osaka University School of Medicine, Kita-ku, Osaka 530, Japan.

[1] The abbreviations used are: L chain, light chain; L_κ and L_λ chain, κ and λ type light chains, respectively; SDS, sodium dodecyl sulfate; FCS, fetal calf serum.

by 1 ml of NCS solubilizer (Amersham) and counted in 10 ml of toluene-based scintillator. For immunoprecipitation of *in vitro* translation products, the reaction mixtures were diluted with 1 ml of 2% Triton X-100, 10 mM L-leucine, 10 mM Tris-HCl, pH 7.4, and 0.15 M NaCl, and centrifuged for 1 h at 100,000 × *g*. The supernatant obtained was mixed with antibodies coupled to Sepharose. ^{125}I-Radioiodination of L_κ chain protein and L_λ chain protein (urinary MOPC41 and RPC20 protein, respectively, Litton Bionetics) was performed as described (18). The specific activity of the iodinated proteins was 2 × 10^6 cpm/μg of protein.

RNA Preparation and Cell-free Protein Synthesis—Cells from 500 ml of culture (1 × 10^6 cells/ml) were harvested, and cytoplasmic RNA was extracted as described previously (19). Wheat germ cell-free system was prepared and used as described (20) with the following modifications: spermidine phosphate was added to 800 μM, and the magnesium concentration was decreased to 1.5 mM (21). [^{3}H]Leucine (60 Ci/mmol) was used for labeling.

SDS-Polyacrylamide Gel Electrophoresis and Fluorography—Proteins and immunoprecipitated Sepharose particles were treated with 0.5 M β-mercaptoethanol and 1.5% sodium dodecyl sulfate at 80°C for 1 min followed by electrophoresis in a 20% polyacrylamide slab gel 2 mm in thickness with a 5% polyacrylamide stacking gel (22). The gel was processed for fluorography as described (23, 24).

Hybridization and Other Methods—RNA excess hybridization was performed at 75°C in 20 mM Tris-HCl, pH 7.4, 0.6 M NaCl, and 0.2 mM ethylenediaminetetraacetic acid. The hybrids formed were assayed by S1 nuclease digestion (6). In all cases, the hybridizable RNA was in at least 8-fold excess over cDNAs used. Thermal stability of cDNA-RNA hybrids was examined by S1 nuclease digestion (6). Oligo(dT)-cellulose chromatography and sucrose density gradient centrifugation were performed as described (6, 25).

RESULTS

All of the 14 Clones of MOPC104E Have Both L_κ and L_λ Chain mRNAs—Fourteen clones of MOPC104E cultured cells were isolated by the microtiter plate technique. L_κ and L_λ chain mRNAs in cytoplasmic RNA of each clone were quantitated by hybridization with cDNA specific to respective mRNA. RNA from clone 4 hybridized with both L_κ and L_λ cDNAs at similar $C_rt_{1/2}$ values around 1.0 (Fig. 1). Since a $C_rt_{1/2}$ value of hybridization is inversely proportional to the concentration of corresponding RNA (26), the results indicate that the RNA contains similar quantities of L_κ and L_λ chain mRNAs. Results of similar experiments with all of the 14 clones are summarized in Table I. The ratio of the amount of L_κ chain mRNA to L_λ chain mRNA ranged from 0.63 to 1.72, with a mean value of 1.02. This is similar to the value of 1.3 obtained for the original MOPC104E tumor cells (10). The mean values for the final extents of hybridization with L_κ and L_λ cDNAs were 56% and 73%, respectively. Purified L_κ and L_λ chain mRNAs hybridized to 72% and 75% with their respective cDNAs. The lower extent of hybridization between L_κ cDNA and MOPC104E RNA may be ascribed to the difference in the variable region sequences, since the L_κ cDNA used is long enough to encode the variable region sequence. The variable region sequences of λ chains cross-hybridize with one another (8).

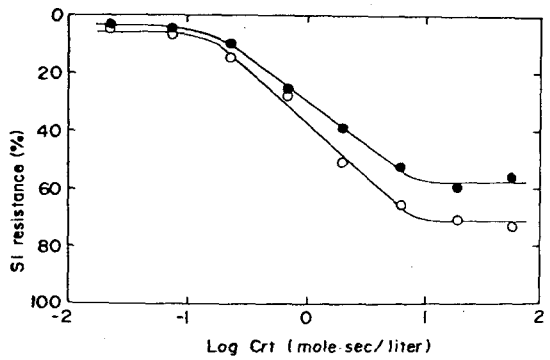


FIG. 1. **Hybridization kinetics of clone 4 RNA to L_κ and L_λ cDNAs.** Cytoplasmic RNA extracted from clone 4 cells (0.136 mg/ml) was hybridized to MOPC41 L_κ cDNA (3.7 × 10^4 cpm/ml) or MOPC104E L_λ cDNA (2.5 × 10^4 cpm/ml). Aliquots were removed at time intervals, and hybrids formed were assayed as described under "Experimental Procedures." ●, L_κ cDNA; ○, L_λ cDNA.

TABLE I

Quantitation of L_κ and L_λ mRNA in RNA of cultured MOPC104E clones

Cytoplasmic RNA of each clone was hybridized with L_κ or L_λ cDNA and $C_rt_{1/2}$ values were determined. The ratios of L_κ to L_λ chain mRNAs were calculated from the ratios of $C_rt_{1/2}$ values for L_λ cDNA to L_κ cDNA, since the $C_rt_{1/2}$ value is inversely proportional to RNA concentration and the genetic complexity of L_κ and L_λ mRNA is identical.

Clone number	$C_rt_{1/2}$		L_κ mRNA/L_λ mRNA
	L_κ cDNA	L_λ cDNA	
1	1.32	1.10	0.83
2	0.63	0.62	0.98
3	1.23	1.30	1.06
4	1.00	1.06	1.06
5	0.80	0.58	0.73
6	0.72	1.24	1.72
7	1.16	1.17	1.01
8	1.25	1.35	1.08
9	0.98	1.39	1.42
10	1.10	1.02	0.93
11	0.95	1.06	1.12
12	1.58	1.00	0.63
13	0.92	0.98	1.07
14	2.40	1.50	0.63

The thermal stability of hybrids formed between L_κ cDNA and MOPC104E RNA was examined by S1 nuclease digestion (Fig. 2). The T_m value (89.2°C) was indistinguishable from that of hybrids formed between L_κ cDNA and MOPC41 L_κ chain mRNA. A similar result was obtained for hybrids formed between MOPC104E RNA and L_λ cDNA.

Characterization of L_κ Chain mRNA in Cloned MOPC104E Cells—The molecular size of L_κ and L_λ chain mRNA in cytoplasmic RNA of clone 7 cells was analyzed by sedimentation in a sucrose gradient under conditions which abolish nonspecific aggregation of RNA (6). Both L_κ and L_λ chain mRNAs sedimented with a single peak of 12.6 S (data not shown), the known size of L chain mRNA (4–9). Approximately 70% of L_κ chain mRNA and 80% of L_λ chain mRNA were adsorbable to oligo(dT)-cellulose, indicating that both types of mRNA have poly(A) sequences. Since the results using RNA from cloned cells shown above were identical to data using RNA from MOPC104E tumor cells (10), subsequent experiments were performed using RNA from uncloned MOPC104E tumor cells.

In Vitro Translation of L_κ and L_λ Chain mRNAs Purified from MOPC104E Tumor—L chain mRNA was purified from MOPC104E tumor by two successive oligo(dT)-cellulose chromatography and sucrose density gradient centrifugations (6). This preparation contained an equal amount of L_κ and L_λ chain mRNAs, each of which comprised about 25% of the total RNA as estimated by hybridization kinetic analysis. Upon electrophoresis in 98% formamide, the RNA migrated as 14 and 18 S bands corresponding to L chain mRNAs and ribosomal RNA, respectively (data not shown). The purified L chain mRNA was translated in the cell-free translation system derived from wheat germ. L_κ and L_λ chain proteins in the translation products were measured by immunoadsorption to rabbit anti-mouse L_κ and L_λ chain antibodies coupled to Sepharose matrices (anti-κ and anti-λ Sepharose), respectively. The immunoadsorption was specific to each type of L chain (Table II). Radioiodinated L_κ and L_λ chain proteins

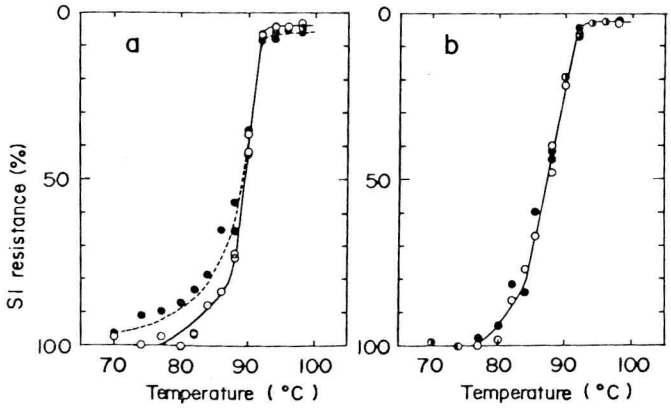


FIG. 2. **Thermal denaturation profiles of hybrids formed between clone 7 RNA and L chain cDNAs.** Cytoplasmic RNA of clone 7 cells was hybridized with L_κ and L_λ cDNAs until a C_0t value of 21 was reached. The hybrids formed were tested for their thermal stabilities in 0.25 M NaCl. Thermal stabilities of hybrids between MOPC41 RNA and L_κ cDNA and between RPC20 RNA and L_λ cDNA were also analyzed as a control. *a*, hybrids with L_κ cDNA: ●, MOPC41 RNA (T_m = 89.0°C); ○, clone 7 RNA (T_m = 89.2°C). *b*, hybrids with L_λ cDNA: ●, RPC20 RNA (T_m = 87.2°C); ○, clone 7 RNA (T_m = 87.2°C).

TABLE II

L chain proteins synthesized in MOPC104E myeloma

L_κ and L_λ chain proteins were determined by immunoadsorption to the saturating amounts of antibody-coupled Sepharoses (300 μl of anti-κ or anti-λ Sepharose suspension/1 ml of reaction mixture) as described under "Experimental Procedures." Radioactivity adsorbed to the ethanolamine-blocked CNBr-activated Sepharose (300 μl) was used as the background value and subtracted. The background values were less than 1.5% of the total radioactivity for *in vivo* proteins and 4% for *in vitro* proteins.

Materials employed	Total protein	Bound to anti-L_κ chain Sepharose	Bound to anti-L_λ chain Sepharose
	cpm	*cpm*	*cpm*
^{125}I-L_κ chain protein	60,000	27,600	120
^{125}I-L_λ chain protein	64,000	240	33,580
In vitro translation product of			
MOPC104E L chain mRNA	38,460	1,730	28,700
MOPC41 L chain mRNA	9,070	4,700	480
In vivo protein			
Extracellular proteins	8,560	60	7,080
Intracellular proteins	30,000	1,050	2,540

were immunoadsorbed as controls. Approximately 75% of the translated material was precipitated with the saturating amount of anti-λ Sepharose. On the other hand, only 4.5% of the translation products bound to the saturating amount of anti-κ Sepharose. Since the mRNA preparation contained an equal amount of L_κ and L_λ chain mRNAs, the L_κ chain mRNA was translated at least 16-fold less efficiently than L_λ chain mRNA in the cell-free translation system. SDS-polyacrylamide gel electrophoresis and fluorography of the material immunoprecipitated by anti-λ Sepharose revealed a single band (24,000 daltons) corresponding to the L_λ chain precursor (Fig. 3, *lane g*) (25, 27, 28). The precipitate with anti-κ Sepharose also gave a faint single band of the same molecular weight (Fig. 3, *lane h*) which might be due to nonspecific adsorption of the L_λ chain precursor to anti-κ Sepharose. As a control, L_κ chain mRNA partially purified from MOPC41 myeloma was translated *in vitro*, and the products were analyzed similarly. About 52% and 5% of the translation products were precipitated with anti-κ and anti-λ Sepharose, respectively.

Absence of L_κ Chain Protein Inside and Outside MOPC104E Cells—In order to determine whether L_κ chain mRNA was translated *in vivo*, we searched for L_κ chain protein inside and outside MOPC104E cells. MOPC104E cells were cultured in the presence of [^{14}C]leucine for 5.5 h. Proteins secreted into the culture fluid were referred to as extracellular proteins, and intracellular proteins were extracted from whole

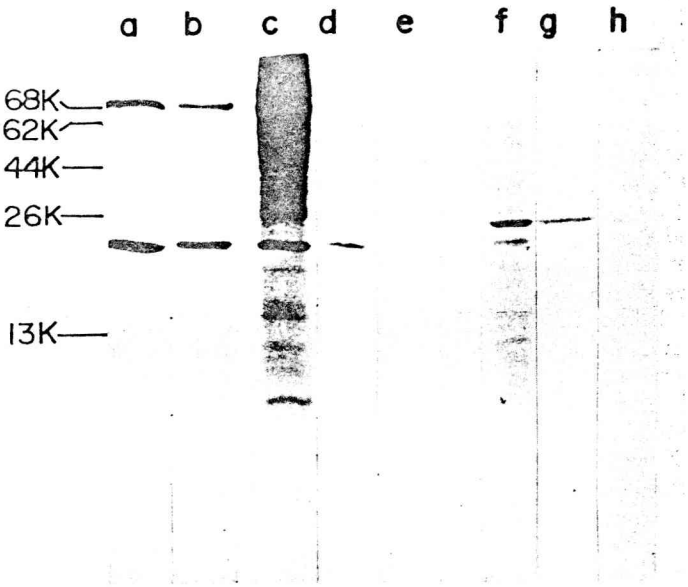


FIG. 3. **SDS-polyacrylamide gel electrophoresis of immunoprecipitated proteins.** Labeled proteins and their immunoprecipitates were treated with 1.3% SDS and 0.5 M β-mercaptoethanol and subjected to polyacrylamide gel electrophoresis. The gel was stained with 0.5% Coomassie Brilliant Blue to stain the molecular weight marker bands and then subjected to fluorography as described under "Experimental Procedures." Molecular weight marker proteins used were bovine serum albumin (68,000 daltons), immunoglobulin γ heavy chain from MOPC31C myeloma (62,000 daltons), chick ovalbumin (44,000 daltons), chymotrypsinogen (26,000 daltons), and equine heart cytochrome *c* (13,400 daltons). *a*, ^{14}C-extracellular proteins (8000 cpm); *b*, ^{14}C-extracellular proteins adsorbed to anti-λ Sepharose (3000 cpm); *c*, ^{14}C-intracellular proteins (23,800 cpm); *d*, ^{14}C-intracellular proteins adsorbed to anti-λ Sepharose (1500 cpm); *e*, ^{14}C-intracellular proteins adsorbed to anti-κ Sepharose (760 cpm); *f*, ^{3}H-*in vitro* translation products of MOPC104E L chain mRNA (20,700 cpm); *g*, ^{3}H-*in vitro* translation products adsorbed to anti-λ Sepharose (7000 cpm); *h*, ^{3}H-*in vitro* translation products adsorbed to anti-κ Sepharose (640 cpm).

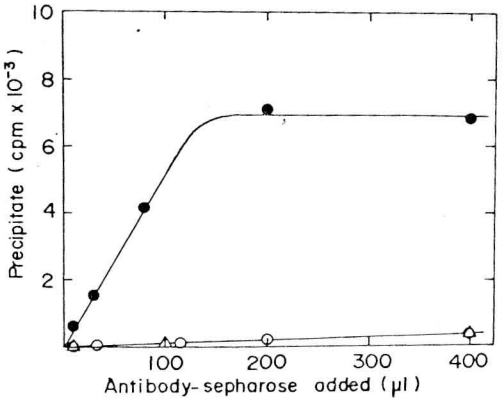


FIG. 4. **Immunoprecipitation of extracellular proteins.** ^{14}C-Labeled extracellular proteins (8600 cpm) were incubated with increasing amounts of anti-λ Sepharose (●), anti-κ Sepharose (○), and ethanolamine-blocked CNBr-activated Sepharose (△). The Sepharoses were treated and counted as described under "Experimental Procedures."

cells with Triton X-100. When extracellular proteins were mixed with increasing amounts of anti-λ Sepharose, radioactivity adsorbed increased linearly and then reached a plateau level (80% of the input radioactivity), whereas essentially no radioactivity was adsorbed to anti-κ Sepharose (Fig. 4 and Table II). The immunoadsorbed material was further analyzed by SDS-polyacrylamide gel electrophoresis. Extracellular proteins bound to anti-λ Sepharose yielded only two bands, corresponding to the molecular weights of immunoglobulin heavy (μ) chain and L chain proteins (Fig. 3, *lane b*). The stronger intensity of L chain band than of heavy chain was explained by the fact that MOPC104E tumor secretes L chain dimer as well as IgM ($\mu_5\lambda_5$) (11).

Intracellular proteins were analyzed similarly (Table II). Approximately 8.5 and 3.5% of the input radioactive protein was adsorbed to anti-λ and anti-κ Sepharose, respectively. SDS-polyacrylamide gel electrophoresis, however, demonstrated the absence of a band corresponding to the L chain in the intracellular proteins adsorbed to anti-κ Sepharose (Fig. 3, *lane e*). Faint bands and smears were seen in the high molecular weight region, suggesting that the adsorbed proteins are not L_κ chain but unknown cross-reactive materials. In contrast, the proteins bound to anti-λ Sepharose contained a prominent L chain band (Fig. 3, *lane d*). These results indicate that MOPC104E myeloma cells synthesize or secrete very little, if any, L_κ chain protein, in confirmation of the reports from other laboratories (11, 29).

DISCUSSION

Presence of Both Types of L Chain mRNAs in MOPC104E Myeloma—It is well established that a single lymphocyte synthesizes only one L chain type and one heavy chain class (1, 2). Furthermore, studies on immunoglobulin allotypic markers have demonstrated that genetic markers on one allele are exclusively expressed in a mature plasma cell of heterozygotic animals (30–32). This phenomenon is called allelic exclusion. The restrictions on types, classes, and allotypic markers of immunoglobulin also hold for myeloma cells (3). MOPC104E is the first case in which an individual myeloma cell has been shown to contain both κ and λ chain mRNAs in an equal quantity. Storb *et al.* (33) reported that MOPC104E and S178 myelomas contain both types of L chain mRNA, in confirmation of our previous report using uncloned myeloma cells (10). Myeloma tumors producing two classes of antibody have been observed (34–36).

L_κ Chain mRNA in MOPC104E is Translated Inefficiently—Although the amount of L_κ chain mRNA is equivalent to that of L_λ chain mRNA in MOPC104E cells, the L_κ chain mRNA is translated very inefficiently in a cell-free system. In addition, we were unable to detect any L_κ chain protein present in the cell or secreted into the culture media in confirmation of the results from other laboratories (11, 29).

On the other hand, L_λ chain mRNA, which is present in the same cell and the same RNA preparation as L_κ chain mRNA, is actively translated *in vivo* as well as *in vitro*, indicating that the translational machinery is working normally in MOPC104E cells. Taken together, these results suggest that MOPC104E L_κ chain mRNA *per se* has an aberrant structure in a region vital to its function. One possibility is the portion close to the 5′ end of mRNA including a cap structure and the ribosomal binding sequence.

Examples of possible defects in mRNA structure have been reported. A variant of mouse myeloma P3 cell line does not produce L chain protein (37). *In vitro* translation of 13 S RNA from these cells does not yield any protein corresponding to L chain precursor, although upon fingerprint analysis the 13 S RNA contains apparently normal ribonuclease T1 oligonucleotides similar to those of L_κ chain mRNA (37). β-Globin mRNA present in a class of °β thalassemia patients was reported not to be translated either *in vivo* or in the cell-free protein synthesis system (38, 39). A recent report on this β-globin mRNA suggested a possible defect in the structure of the mRNA molecule itself (40).

L Chain Genes in MOPC104E Cells—Recent studies on immunoglobulin genes indicate that the gene conversion plays an important role in the expression of immunoglobulin genes in lymphocyte (41–45). Brack *et al.* (41) have clearly demonstrated that the variable and constant region genes which are far apart in embryonic tissues are brought together in close proximity in plasma cells expressing the particular gene. There seems to be a mechanism which restricts the gene conversion to one type of L chain genes, since only the gene for that type of the L chain which is expressed in plasmacytomas has undergone the recombination (41, 42). In this regard, it would be interesting to know whether or not L_κ chain mRNA of MOPC104E contains a variable region sequence.

Several possibilities may account for synthesis of two types of L chain mRNA in MOPC104E cells. These include chromosomal abnormalities often found in myeloma cells (46), fusion of different myeloma cells or lymphocytes (47, 29), and unknown defects in the regulation of the specific gene activation. The phenomenon reported here has two abnormalities: (*a*) transcription of the L_κ chain gene, which is normally repressed; and (*b*) the dysfunction of this L_κ chain mRNA. These results could be explained by two unusual events in the MOPC104E myeloma. However, we prefer to hypothesize that synthesis of the inactive L_κ chain mRNA may be attributable to an aberrant conversion of the gene. The structure of the L chain gene in MOPC104E may give some insights into the mechanism of the gene conversion and its relationship to the expression of immunoglobulin genes.

REFERENCES

1. Bernier, G. M., and Cebra, J. J. (1964) Science **144**, 1590–1591
2. Burtin, P., and Buffe, D. (1963) *Proc. Soc. Exp. Biol. Med.* **114**, 171–175
3. Potter, M. (1972) *Physiol. Rev.* **52**, 632–719
4. Stavnezer, J., Huang, R. C. C., Stavnezer, E., and Bishop, J. M. (1974) *J. Mol. Biol.* **88**, 43–63.
5. Faust, C. H., Diggelman, H., and Mach, B. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71**, 2491–2495
6. Honjo, T., Packman, S., Swan, D., Nau, M., and Leder, P. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3659–3663
7. Tonegawa, S., Steinberg, C., Dube, S., and Bernardini, A. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71**, 4027–4031
8. Honjo, T., Swan, D., Nau, M., Norman, B., Packman, S., Polsky, F., and Leder, P. (1976) *Biochemistry* **15**, 2775–2779
9. Tonegawa, S. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 203–207
10. Ono, M., Kawakami, M., Kataoka, T., and Honjo, T. (1977) *Biochem. Biophys. Res. Commun.* **74**, 796–802
11. McIntire, K. R., Asofsky, R. M., Potter, M., and Kuff, E. L. (1965) *Science* **150**, 361–362
12. Fahey, J. L., McCoy, P. F., and Goulian, M. (1958) *J. Clin. Invest.* **37**, 272–284
13. Studier, F. W. (1965) *J. Mol. Biol.* **11**, 373–390
14. Britten, R. J., Graham, D. E., and Neufeld, B. R. (1974) *Methods Enzymol.* **29**, 363–418
15. Axen, R., Porath, J., and Ernback, S. (1967) *Nature* **214**, 1302–1304
16. Cuatrecasas, P. (1971) *Methods Enzymol.* **22**, 345–378
17. Rhoads, R. E., McKnight, G. S., and Schimke, R. T. (1971) *J. Biol. Chem.* **246**, 7407–7410
18. Hunter, W. M., and Greenwood, F. C. (1962) *Nature* **194**, 495–496
19. Honjo, T., Shimizu, A., Tsuda, M., Natori, S., Kataoka, T., Dohmoto, C., and Mano, Y. (1977) *Biochemistry* **26**, 5764–5769
20. Roberts, B. E., and Paterson, B. M. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **70**, 2330–2334
21. Atkins, J. F., Lewis, J. B., Anderson, C. W., and Gesteland, R. F. (1975) *J. Biol. Chem.* **250**, 5688–5695

22. Blattler, D. P., Gardner, F., Van Slyke, K., and Bradley, A. (1972) *J. Chromatogr.* **64,** 147–155
23. Bonner, W. M., and Laskey, R. A. (1974) *Eur. J. Biochem.* **46,** 83–88
24. Laskey, R. A., and Mills, A. D. (1975) *Eur. J. Biochem.* **56,** 335–341
25. Swan, D., Aviv, H., and Leder, P. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69,** 1967–1971
26. Bishop, J. O. (1969) *Biochem. J.* **113,** 805–811
27. Milstein, C., Brownlee, G. G., Harrison, T. M., and Mathews, M. B., (1972) *Nature New Biol.* **239,** 117–120
28. Burstein, Y., Zemell, R., Kantor, F., and Schechter, I. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 3157–3161
29. Margulies, D. H., Kuehl, W. H., and Scharff, M. D. (1976) *Cell* **8,** 405–415
30. Pernis, B. G., Chiappino, G., Kelus, A. S., and Gell, B. G. H. (1965) *J. Exp. Med.* **122,** 853–876
31. Cebra, J. J., Colberg, J. E., and Dray, S. (1966) *J. Exp. Med.* **123,** 547–558
32. Mäkelä, O., and Cross, A. W. (1970) *Prog. Allergy* **14,** 145–207
33. Storb, U., Hager, L., Wilson, R., and Putnam, D. (1977) *Biochemistry* **16,** 5432–5438
34. Takahashi, M., Takagi, N., Yagi, Y., Moore, G. E., and Pressman, D. (1969) *J. Immunol.* **102,** 1388–1393
35. Litwin, S. D., Lin, P. K., Huetterroth, T. H., and Cleve, H. (1973) *Nature New Biol.* **246,** 179–181
36. Hopper, J. E. (1975) *J. Exp. Med.* **115,** 1101–1107
37. Cowan, N. J., Secher, D. S., and Milstein, C. (1974) *J. Mol. Biol.* **90,** 691–701
38. Kan, Y. W., Holland, S. P., Dozy, A. M., and Varmus, H. E. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72,** 5140–5144
39. Temple, G. F., Chang, J., and Kan, Y. W. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 3047–3051
40. Old, J. M., Proudfoot, N. J., Wood, W. G., Longley, J. I., Clegg, J. B., and Weatherall, D. J. (1978) *Cell* **14,** 289–298
41. Brack, C., Hirama, M., Lenhard-Schuller, R., and Tonegawa, S. (1978) *Cell* **15,** 1–14
42. Tonegawa, S., Hozumi, N., Matthyssens, G., and Schuller, R. (1976) *Cold Spring Harbor Symp. Quant. Biol.* **41,** 877–889
43. Sakano, H., Hüppi, K., Heinrich, G., and Tonegawa, S. (1979) *Nature* **280,** 288–294
44. Honjo, T., and Kataoka, T. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75,** 2140–2144
45. Max, E. E., Seidman, J. G., and Leder, P. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 3450–3454
46. Yoshida, T. H., Imai, H. T., and Potter, M. (1968) *J. Natl. Cancer Inst.* **41,** 1083–1095
47. Köhler, G., and Milstein, C. (1975) *Nature* **256,** 495–497

Reprinted from
Proc. Natl. Acad. Sci. USA
Vol. 76, No. 9, pp. 4240–4244, September 1979
Biochemistry

参考論文

Cloning immunoglobulin γ2b chain gene of mouse: Characterization and partial sequence determination

(constant region gene/R-loop mapping/DNA sequence/intervening sequence)

TOHRU KATAOKA*, YURIKO YAMAWAKI-KATAOKA*, HIDEO YAMAGISHI†, AND TASUKU HONJO*

*Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo; and †Department of Biophysics, Faculty of Science, Kyoto University, Kyoto, Japan

Communicated by Osamu Hayaishi, May 14, 1979

ABSTRACT DNA from newborn mice was digested with restriction endonuclease *Eco*RI, and a 6.6-kilobase fragment encoding immunoglobulin γ2b chain mRNA derived from MPC 11 myeloma was enriched about 100-fold by RPC-5 column chromatography and agarose gel electrophoresis. The 6.6-kilobase fragment was cloned with λgt *WES*·λB as EK2 vector. The cloned phage (λgt *WES*·IgH22) contained the constant region gene of the γ2b chain but not the variable region gene of MPC 11 mRNA. The constant region genes of the other γ chains (i.e., γ1, γ2a, and γ3) were not present in λgt *WES*·IgH22 DNA. R-loop mapping indicates that the γ2b chain structural gene is divided into two parts (330 ± 60 SD base pairs and 930 ± 110 SD base pairs) by an intervening sequence (360 ± 100 SD base pairs). The nucleotide sequence around the junction of the hinge region and C_H2 domain was determined and shown to match the amino acid sequence of the initial part of the C_H2 domain of the γ2b chain. The base sequence upstream from the junction, however, is unrelated to the amino acid sequence of the C_H1 domain and the hinge region of all the γ chains whose sequences have been determined. These results indicate that the γ2b chain gene is interrupted at the junction of the hinge region and C_H2 domain by an intervening sequence. The existence of two more intervening sequences, one between the C_H1 domain and the hinge region and the other between the C_H2 and C_H3 domains, is discussed.

Immunoglobulin heavy chain genes consist of a family of variable region (V) genes and several constant region (C) genes. V and C genes are genetically linked and expressed as a *cis* combination (1–3). In a given lymphocyte, immunoglobulin genes on one allele are exclusively expressed (4–6). Recently, we have shown that specific C genes are deleted from one allele in mouse myelomas depending on which C genes are expressed (7, 8). We have proposed an allelic deletion model as the mechanism of the V–C gene recombination that results in the expression of the specific immunoglobulin genes (7, 8). According to the model, a chromosomal segment that is located between recombining V and C genes is excised out from one allele so as to bring C and V genes together. A linear arrangement of heavy chain C genes is also proposed as follows: V genes, spacer, μ, γ3, γ1, γ2b, γ2a, and α.

The use of recombinant DNA technology to clone a given segment of the mammalian genome offers the most direct means of testing the allelic deletion model. With this goal in mind, we have set out to clone immunoglobulin heavy chain genes from mouse genome. Recent advances in the availability of a vector (9) and in purification of the mammalian gene (10) have enabled us to clone the immunoglobulin heavy chain gene (γ1 C gene) in EK2 phage λ, λgt*WES*·λB (11). As the first step in testing the model, we wish to determine the relative location of γ chain genes by cloning a large piece of DNA that contains two or more neighboring genes. For this purpose it is essential to clone two putative adjacent γ chain genes. In this paper we will describe cloning of the γ2b chain C gene, which is proposed to be adjacent to the γ1 chain gene, and its structure.

EXPERIMENTAL PROCEDURES

Partial Purification of DNA Fragment. High molecular weight DNA was extracted from whole newborn mice (killed within 24 hr after birth). DNA was digested with *Eco*RI, extracted with phenol, and precipitated with ethanol. Forty milligrams of *Eco*RI-digested DNA was chromatographed on an RPC-5 column (0.8 × 90 cm) as described (10, 12). Fractions that hybridized with the γ2b chain cDNA by *in situ* hybridization were pooled and concentrated by ethanol precipitation. The DNA obtained was applied to a 1% agarose gel (7 × 15 × 4 cm) and fractionated as described (13). Fractions were assayed by *in situ* hybridization with nick-translated γ2b chain [^{32}P]cDNA cloned in a plasmid (pG2b-4).

Preparation and Screening of Recombinant Phage. Bacteriophage λgt*WES*·λB (9) was used as the EK2 vector and propagated in ED8656 (14). Cloning experiments were carried out in a P2 facility (15, 16). Bacterial strains NS428 [N205 (λ Aam11 b2 red 3 *cI*857 Sam7)] (17) and λdg805 [W3350 (λdgal805 *cI*857 Sam7)] were used for *in vitro* packaging.

Partially purified 6.6-kilobase (kb) mouse DNA (2 μg) was ligated with 5.1 μg of λgt*WES* outer fragments in a 50-μl reaction mixture with 0.6 unit of T4 ligase as described (18). The recombinant DNA was packed into coat proteins *in vitro* and the recombinant phages were plated by the method of Blattner *et al.* (19). Twenty sets of about 2000 plaques were directly transferred from LB broth agar plates to Millipore filters as described (20). The Millipore filters were hybridized to ^{32}P-labeled *Hha* I fragment of pG2b-4 and autoradiographed. Recombinant phage DNA was prepared as described (18).

***In Situ* Hybridization.** DNA, fixed on a Millipore filter (21), was hybridized to an appropriate probe as described by Jeffrey and Flavell (22) except that 1 M NaCl/50 mM Tris·HCl, pH 7.4/10 mM EDTA replaced 0.9 M NaCl/0.09 M sodium citrate. The probes used were [^{32}P]cDNA (specific activity, 50 cpm/pg) complementary to MPC 11 mRNA (7, 23, 24) or a ^{32}P-labeled *Hha* I fragment (specific activity, 60 to ≈100 cpm/pg) of pG2b-4 DNA. When [^{32}P]cDNA was used as a probe, oligo(dT)-cellulose-bound mRNA from MOPC 511 myeloma (IgA producer) was added to 10 μg/ml. The plasmid pG2b-4 is a hybrid of pBR322 (25) and an ≈250-base sequence of γ2b chain cDNA (unpublished data). The γ2b chain insert was isolated as the *Hha* I fragment of 460 base pairs in length. Nick translation of DNA was done as described, with [α-^{32}P]dCTP (300 to ≈400 Ci/mmol, 1 Ci = 3.7 × 10^{10} becquerels) (New England Nuclear) (26).

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

Abbreviations: V and C, variable and constant region, respectively; kb, kilobase.

Hybridization in Solution. Phage or plasmid DNA was sonicated, heat-denatured, and hybridized to the γ chain cDNA as described (27). Hybrid formed was determined by S1 nuclease resistance (27). cDNAs complementary to the purified γ chain mRNAs were synthesized as described (23). Specific activity of [^{3}H]cDNA was 17 cpm/pg. The γ chain mRNA used in the hybridization reaction was partially purified from mouse myeloma by an oligo(dT)-cellulose column as described (28). The purity of the mRNA was about 4%.

Electron Microscopy. R-loops (29, 30) between the cloned 6.6-kb insert DNA and MPC 11 γ2b chain mRNA were formed with 8.7 μg of DNA per ml and 16 μg of RNA per ml in 70% (vol/vol) formamide 0.1 M Tricine (*N*-[tris(hydroxymethyl)-methyl]glycine)-NaOH, pH 8.0/0.5 M NaCl/10 mM EDTA at 56°C for 12 hr. The sample was diluted, spread, stained, and shadowed as described (31).

DNA Sequence Analysis. The 6.6-kb insert DNA of λgt*WES*·IgH22 was separated from λgt*WES* outer fragments by *Eco*RI digestion and subsequent preparative agarose gel electrophoresis. The DNA (200 μg) was cleaved with *Sac* I, and the 520-base fragment was isolated on 3% polyacrylamide gels, treated with bacterial alkaline phosphatase, and labeled at the 5′ ends with polynucleotide kinase. The DNA was cleaved with *Hap* II and the 280-base fragment was isolated on 5% polyacrylamide gels. The fragment was chemically modified and analyzed as described (32).

Enzymes. *Eco*RI (33), *Sac* I (R. J. Roberts, personal communication), *Hap* II (34), and T4 polynucleotide kinase (35) were purified as described. T4 ligase was a generous gift of M. Takanami (Kyoto University). Other restriction endonucleases were obtained from New England BioLabs.

RESULTS

Partial Purification of the Immunoglobulin γ2b Chain Gene. In order to enrich the immunoglobulin gene sequence, we used sequential purification steps of RPC-5 column chromatography and agarose gel electrophoresis that were successfully adopted to cloning of the globin gene (10). DNA derived from whole newborn mice was digested with restriction endonuclease *Eco*RI. Approximately 40 mg of *Eco*RI-digested DNA was fractionated by RPC-5 chromatography. Aliquots of pooled fractions were electrophoresed in an agarose gel (36, 37) and DNA was transferred to a Millipore filter according to Southern (21). The Southern blot was hybridized with MPC 11 (γ2b) cDNA and autoradiographed. Fractions containing a marked 6.6-kb band were pooled, and approximately 3.5 mg of DNA obtained was applied to a preparative agarose gel. Aliquots of DNA fractions were again electrophoresed in a 1% agarose gel and blotted to a Millipore filter. The filter was hybridized with MPC 11 cDNA cloned in pBR322. As shown in Fig. 1, the 6.6-kb fragment was found in fractions 14 and 15. No other bands were seen in all the fractions obtained (up to 20 kb). Fraction 14 was concentrated by ethanol precipitation, and approximately 0.1 mg of DNA was obtained.

Isolation and Identification of a Recombinant Phage Containing the Immunoglobulin γ2b Chain Gene. Purified outer fragments of λgt*WES*·λB were annealed with the DNA from fraction 14, ligated, and used for *in vitro* packaging. After screening about 4×10^4 phages we found four recombinants containing the γ2b gene sequence. One of them was referred to as λgt*WES*·IgH22 and grown in large quantity. DNA was extracted and hybridized with MPC 11 (γ2b) cDNA in solution (Table 1). γ2b cDNA hybridized with λgt*WES*·IgH22 but not with λgt*WES*·λB. The other three recombinants were similar to λgt*WES*·IgH22 in the extent of hybridization to γ2b cDNA and the cleavage sites of several restriction endonucleases.

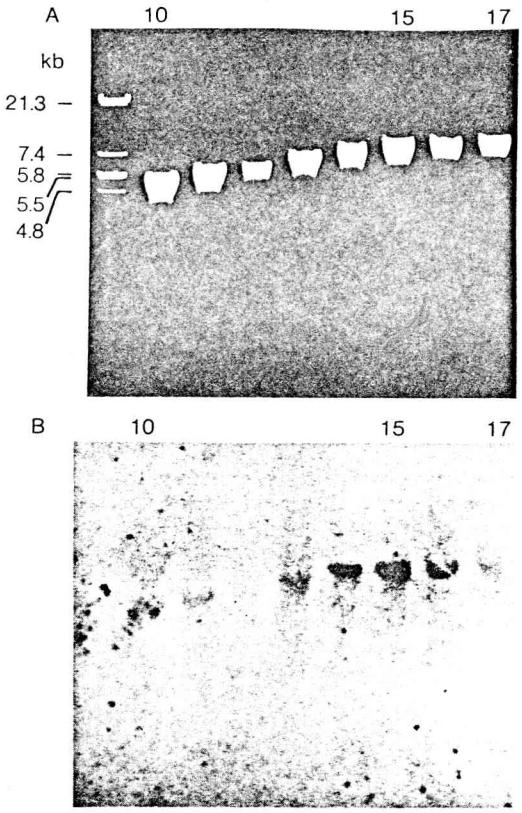


FIG. 1. Preparative agarose gel electrophoresis of *Eco*RI-digested mouse DNA. (*A*) Ethidium bromide-stained agarose gel of preparative agarose gel fractions. Migration is from top to bottom. Only the relevant eight fractions are presented here. (*B*) Autoradiogram of the Southern blot of the gel hybridized with pG2b-4 [^{32}P]DNA. Fraction numbers are shown on the top of each lane. *Eco*RI-cleaved λ*cI*857 was used as size markers.

Substantially less λgt*WES*·IgH22 DNA hybridized with MPC 11 cDNA than with MPC 11 mRNA and MOPC141 mRNA containing the same C sequence and a V sequence different from that of MPC 11 mRNA (Table 1). The extent of hybridization did not increase even when the amount of λgt*WES*·IgH22 DNA was increased by 10-fold. As a control, plasmid pG2b-4, containing about 250-base γ2b cDNA, hybridized to 19.6% of the cDNA in proportion to the length of the inserted cDNA. The [^{3}H]cDNA probe used was approximately 1800 bases long and estimated to encode the whole C sequence, including the 3′-untranslated portion, but only a part, if any, of the V sequence (38). These results indicate that λgt*WES*·IgH22 contains a considerable part of the γ2b chain gene although it may not encompass the whole sequence.

Table 1. Hybridization of λgt*WES*·IgH22 DNA with MPC11 γ2b cDNA

DNA or mRNA	% hybridization with MPC11 γ2b cDNA
λgt*WES*·IgH22	67.9
λgt*WES*·λB	0.6
pG2b-4	19.6
MOPC141 mRNA	84.3
MPC11 mRNA	89.5

Hybridizations of DNA (3 μg for λ phage DNA, 2 μg for pG2b-4 DNA) or mRNA (3.8 μg) were performed in the presence of 2100 cpm of MPC11 [^{3}H]cDNA in 50 μl of reaction mixture at 65°C for 2 hr. The hybrids formed were assayed by S1 nuclease digestion.

Table 2. Hybridization of λgt*WES*·IgH22 DNA with γ chain cDNAs

cDNA	% hybridization
MOPC31C (γ1)	9.6
HOPC1 (γ2a)	17.9
MPC11 (γ2b)	75.7
J606 (γ3)	11.2

Hybridization of λgt*WES*·IgH22 DNA (3 μg) was performed in the presence of about 2000 cpm of [^{3}H]cDNA. The extent of hybridization of each cDNA was normalized, taking the extent of hybridization with homologous mRNA as 100%.

In order to test whether λgt*WES*·IgH22 contains the genes of the other γ subclasses, we hybridized λgt*WES*·IgH22 DNA with the other γ chain cDNAs. Table 2 shows that λgt*WES*·IgH22 hybridized specifically with γ2b cDNA. Small but significant extents of hybridization with γ1, γ2a, and γ3 cDNAs seem to be attributable to partial homology shared among the γ chain C gene sequences (38).

Location of Structure Sequence. Both λgt*WES*·IgH22 DNA and the 6.6-kb insert DNA were digested with several restriction endonucleases and electrophoresed in 1.2% agarose gel. The restriction map (see Fig. 4) was constructed by a comparison of the size of a new insert fragments with known adjacent enzyme sites in λgt*WES* outer fragments and by combined cleavage with several restriction endonucleases (data not shown). In order to localize the γ2b chain gene in the 6.6-kb insert DNA, the DNA fragments were transferred to nitrocellulose filters and hybridized with MPC 11 [^{32}P]cDNA synthesized with random oligonucleotides as primer (24). As shown in Fig. 2, *Sac* I cleaved the 6.6-kb insert DNA into five fragments of 3.8, 2.1, 0.52, 0.3, and 0.28 kb. The five fragments were placed into order as shown in Fig. 4. The fragments of 2.1 and 0.52 kb hybridized with MPC 11 cDNA. The fragments of 3.8 and 0.28 kb hybridized weakly and their bands on autoradiography were seen only by longer exposure (data not shown). *Bam*HI cut the 6.6-kb insert into 4.9-, 1.3-, 0.25-, 0.17-, and 0.09-kb fragments (see Fig. 4). The fragment next to the left-end fragment (1.3 kb) hybridized very weakly with MPC 11 cDNA; the 4.9- and 0.25-kb fragments hybridized intensely. Among three fragments (3.0, 2.8, and 0.84 kb) produced by digestion of the 6.6-kb insert DNA with *Kpn* I, the middle fragment (0.84 kb) did not hybridize with the cDNA. These results suggest that the γ2b chain sequence is present somewhere between 1.3 and 3.0 kb from the left end of the insert DNA. *Xho* I seems to cut within the γ2b chain sequence.

R-Loop Mapping. To further analyze the structure of the γ2b chain gene, R-loops were formed between the 6.6-kb insert DNA and MCP 11 γ2b chain mRNA (Fig. 3). Two R-loops of unequal size virtually adjacent to one another were interrupted by a looped-out, double-stranded region. Such a structure indicates the presence of an intervening sequence within the structure sequence for the γ2b chain, as has been reported in other eukaryote genes (30, 39, 40). A short whisker of mRNA was observed at the outer end of the smaller R-loop. A series of R-loops formed between the 6.6-kb insert DNA and MPC 11 mRNA were measured with the 6.6-kb insert as a duplex standard. The mean lengths (±SD) of the shorter R-loop, the looped-out region, and the longer R-loop were estimated to be 330 ± 60, 360 ± 100, and 930 ± 110 base pairs, respectively. The left and right double-stranded regions were 1220 ± 75 and 3770 ± 160 base pairs long, respectively. The results are consistent with the conclusion based on the restriction mapping that the γ2b chain sequence is present somewhere between 1.3 and 3.0 kb from the left end of the insert DNA.

Base Sequence Determination around Junction of Intervening and Coding Sequences. Our interpretation of R-loops and restriction site analyses could be directly confirmed by the determination of the nucleotide sequence around the junction of the intervening and coding sequences. A *Sac* I site 2100 base pairs away from the left end of the insert was expected to be close to the junction of the larger R-loop and looped-out region. The base sequence starting from this site to the right was determined by the chemical modification method of Maxam and Gilbert (32). As shown in Fig. 4, the nucleotide sequence matched the known amino acid sequence perfectly at positions 236–259 of mouse γ2b chain (41), except for position 241, with the direction of the gene from left to right. The codon for po-

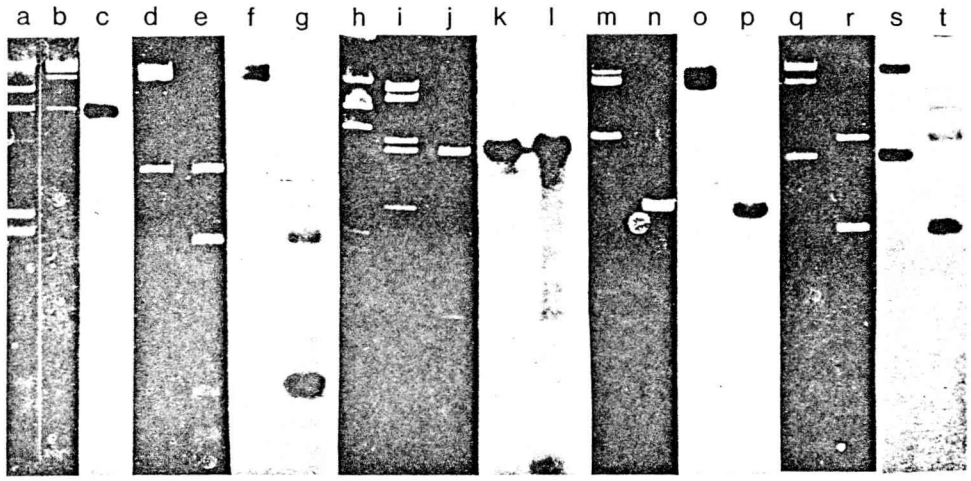


FIG. 2. Digestion of λgt*WES*·IgH22 DNA and its insert DNA by various restriction endonucleases. Approximately 1.6 μg of λgt*WES*·IgH22 DNA or 0.8 μg of the insert DNA was cleaved by the restriction endonucleases indicated and electrophoresed in 1.2% agarose gels. The DNA fragments were transferred to a nitrocellulose filter and hybridized with γ2b [^{32}P]cDNA synthesized with random oligonucleotides as primers (24). Size markers used were from a *Hin*dIII digest of λ*cI*857 DNA (lanes a and h). The sizes of the marker fragments were 23.5, 9.8, 6.6, 4.5, 2.5, 2.2, and 0.6 kb from top to bottom. The following pairs of figures are ethidium bromide stains and autoradiograms of Southern blot for cleavage of λgt*WES*·IgH22 DNA by *Eco*RI (lanes b and c), *Sac* I (lanes d and f), *Bam*HI (lanes i and k), *Kpn* I (lanes m and o), and *Xho* I (lanes q and s) and of the insert DNA by *Sac* I (lanes e and g), *Bam* HI (lanes j and l), *Kpn* I (lanes n and p), and *Xho* I (lanes r and t). The location of small outer fragments were determined by cleavage of terminally ^{32}P-labeled 6.6-kb insert DNA.

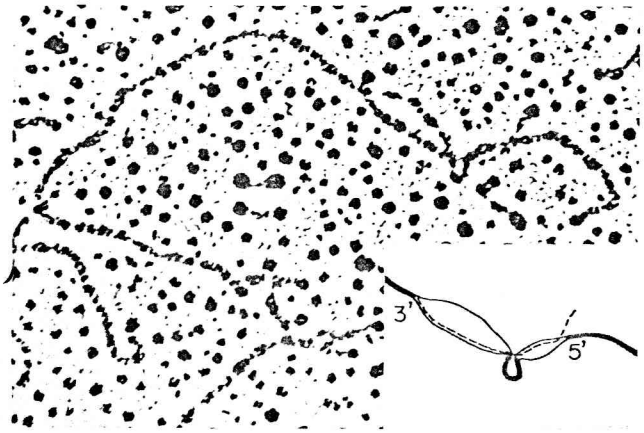


FIG. 3. Electron micrograph of R-loops between the 6.6-kb insert DNA of λgt*WES*·IgH22 DNA and MPC 11 mRNA. R-loops formed were visualized by a magnification of 60,000. (*Inset*) Diagram of the structure. Heavy lines represent double-stranded DNA, narrow lines represent single-stranded DNA, and broken lines indicate γ2b chain mRNA.

sition 241 predicts the presence of Ser, which coincides with the γ2a chain but not with the γ2b chain. The discrepancy is likely to be an error in amino acid sequence. The base sequence predicts the amino acid sequence of Leu-Ser-Ser-Ser-Ala-Pro-Asn for positions 229–235, although the amino acid sequence data are unavailable for the NH_2 terminus to position 236 of the γ2b chain. Because the known amino acid sequence of the γ2b chain shares extensive homology with that of the γ2a chain, we have compared the predicted sequence with the amino acid sequence of the γ2a chain. As shown in Fig. 4, the homology extends to Ala-Pro-Asn (positions 233–235), which is considered to be the beginning of the C_H2 domain (41, 42). Upstream from this junction, however, essentially no homology

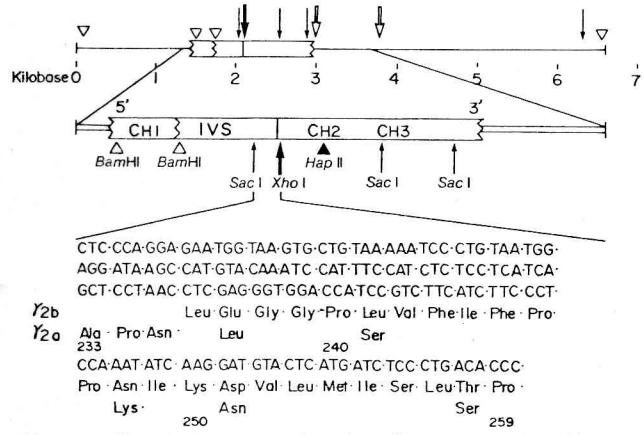


FIG. 4. Restriction map and nucleotide sequence of 6.6-kb insert DNA of λgt*WES*·IgH22. The top line indicates the position of γ2b chain gene within the 6.6-kb insert DNA. The direction of transcription is from left (5′) to right (3′). The entire γ2b chain gene estimated by R-loop and restriction site mapping is represented by the boxes. The middle line shows an enlargement of the gene and its neighborhood. An intervening sequence is labeled as IVS. The serrated lines in the box indicate that the borders are not definitely assigned. The restriction endonuclease cleavage sites are indicated by (↓) *Sac* I, (⬇) *Xho* I, (⬇) *Kpn* I, (▽) *Bam*HI, and (▼) *Hap* II. The bottom lines are the nucleotide sequences of the anti-coding strand starting from the *Sac* I site towards the right. The *Xho* I site is indicated by a box. The amino acid sequences of the γ2b and γ2a chains are shown below the nucleotide sequence. Only those amino acid residues of the γ2a chain that are different from the corresponding residues of the γ2b chain are indicated.

was seen between two sequences. The cysteine and proline residues at positions 231 and 232 is the COOH terminus of the hinge region and is highly conserved among mouse γ1 (43), mouse γ2a (42), guinea pig γ2 (44, 45), rabbit γ (46, 47), and human Eu γ1 (48) chains although the base sequence predicts Ser-Ser for these positions. There is a long stretch of the conserved amino acid sequence (Thr-Lys-Val-Asp-Lys) at positions 210–215, which does not coincide with the base sequence. Furthermore, a termination codon (TAA) was found 46 to ≈48 base pairs upstream from the junction in the same frame with the matched sequence. These results tempted us to conclude that an intervening sequence splits the γ2b chain gene somewhere very close to the junction of the hinge region and the C_H2 domain.

DISCUSSION

The 6.6-kb DNA fragment isolated from newborn mice was shown unambiguously to contain the γ2b chain C gene and its flanking sequence. The results of the R-loop study and the nucleotide sequence determination show that it is likely that the smaller R-loop region (330 ± 60 base pairs) codes for the C_H1 domain and the larger R-loop region (930 ± 110 base pairs) codes for the C_H2 and C_H3 domains. The looped-out, double-stranded region seems to correspond to the intervening sequence that interrupts the γ2b chain gene at the junction of the hinge region and the C_H2 domain. Recently we have completed a base sequence determination of the γ2b chain gene. The results demonstrate that intervening sequences separate the hinge region from the C_H1 domain and the C_H2 domain from the C_H3 domain. Less than ≈30% of the molecules observed by electron microscopy have the structure indicating the presence of a short intervening sequence at the middle of the large R-loop, suggesting that our conditions for formation of R-loops were not optimal for detecting a short intervening sequence (about 100 bases). The smaller whisker at the outer end of the smaller R-loop probably corresponds to the V sequence of MPC 11 mRNA, which is not encoded by λgt*WES*·IgH22.

There are several myeloma variants that produce heavy chain proteins of altered structure. Among them the most common is the deletion of a whole domain from the C sequence, which comprises a tandem array of three or four domains (49–53). A mouse γ1 chain mutant (IF2) was also reported to lack the whole C_H1 domain (43). Another set of variants seems to arise by recombination between different C genes at the junction of domains (54, 55). These phenomena tempted each domain to be conceived of as a separate gene (50). The idea seems to be valid because each domain is separated from the other. It will be interesting to see whether the presence of the intervening sequence facilitates the recombination rate or whether improper splicing is responsible for these abnormalities.

The intervening sequence and its splicing mechanism have been proposed to play an important role in the evolution of eukaryote genes by linking duplicated genes or prototype peptide genes that are not directly adjacent to each other (56, 57). The hypothesis is consistent with our results that intervening sequences split immunoglobulin heavy chain genes at the junction of functional units of the protein (i.e., domains and the hinge region). A phylogenic study based on the comparison of amino acid sequences indicates that the heavy chain protein has evolved by successive duplication of a domain (48). Because precisely contiguous duplication of a domain may be a rather rare event, introduction of the splicing mechanism may have facilitated construction of a new gene by linking "unsuccessfully" duplicated domains. After submission of this manuscript, Sakano *et al.* (58) and Early *et al.* (59) also reported that γ1 and α chain genes are interrupted by intervening sequences at the junction of domains and the hinge region.

We thank Dr. Y. Mano for his continuous encouragement. We also thank Dr. P. Leder, Dr. F. Blattner, and Dr. R. J. Roberts for bacterial strains, phages, and unpublished protocols. Thanks are also extended to Drs. M. Kawakami and M. Ono for preparation of purified H chain mRNAs. This investigation was supported in part by grants from the Matsunaga Science Foundation and the Ministry of Education, Science and Culture of Japan.

1. Kindt, T. J., Mandy, W. J. & Todd, C. W. (1970) *Biochemistry* **9**, 2028–2032.
2. Landucci-Tosi, S., Mage, R. G. & Dubiski, S. (1970) *J. Immunol.* **104**, 641–647.
3. Knight, K. L. & Hanly, W. C. (1975) in *Contemporary Topics in Molecular Immunology*, eds. Inman, F. P. & Mandy, W. J. (Plenum, New York), Vol. 4, pp. 55–88.
4. Pernis, B., Chiappino, G., Kelus, A. S. & Gell, P. G. H. (1965) *J. Exp. Med.* **122**, 853–876.
5. Cebra, J. J., Colberg, J. E. & Dray, S. (1966) *J. Exp. Med.* **123**, 547–558.
6. Mäkelä, O. & Cross, A. W. (1970) *Prog. Allergy* **14**, 145–207.
7. Honjo, T. & Kataoka, T. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2140–2144.
8. Honjo, T. & Kataoka, T. (1979) in *Oncogenic Viruses and Host Cell Genes*, eds. Ikawa, Y. & Odaka, T. (Academic, New York), pp. 461–473.
9. Leder, P., Tiemeier, D. & Enquist, L. (1977) *Science* **196**, 175–177.
10. 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.
11. Honjo, T. (1978) in *Immunology '78*, eds. Yamamura, Y. & Tada, T. (Nakayama-Shoten, Tokyo), pp. 75–82.
12. Pearson, R. L., Weiss, J. F. & Kelmers, A. D. (1971) *Biochim. Biophys. Acta* **228**, 770–774.
13. Polsky, F., Edgell, M. H., Seidman, J. G. & Leder, P. (1978) *Anal. Biochem.* **87**, 397–410.
14. Wood, W. B. (1966) *J. Mol. Biol.* **16**, 118–133.
15. Recombinant DNA Research Guidelines (1978) *Jpn. Sci. Monthly* **31**, 626–647.
16. Recombinant DNA Research Guidelines (1976) *Fed. Reg.* (Sept. 9, 1976).
17. Sternberg, N., Tiemeier, D. & Enquist, L. (1977) *Gene* **1**, 255–280.
18. Tiemeier, D. C., Tilghman, S. M. & Leder, P. (1977) *Gene* **2**, 173–191.
19. Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Richards, J. E., Slighton, J. L., Tucker, P. W. & Smithies, O. (1978) *Science* **202**, 1279–1284.
20. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
21. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
22. Jeffrey, A. J. & Flavell, R. A. (1977) *Cell* **12**, 429–439.
23. Ono, M., Kondo, T., Kawakami, M. & Honjo, T. (1977) *J. Biochem. (Tokyo)* **81**, 949–954.
24. Taylor, J. M., Illmensee, R. & Summers, J. (1976) *Biochim. Biophys. Acta* **442**, 324–330.
25. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heynecker, H. L. & Boyer, H. W. (1977) *Gene* **2**, 95–113.
26. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184–1188.
27. Honjo, T., Packman, S., Swan, D., Nau, M. & Leder, P. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3659–3663.
28. Honjo, T., Packman, S., Swan, D. & Leder, P. (1976) *Biochemistry* **15**, 2780–2785.
29. Thomas, M., White, R. L. & Davis, R. W. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2294–2298.
30. White, R. L. & Hogness, D. S. (1977) *Cell* **10**, 177–192.
31. Yamagishi, H., Inokuchi H. & Ozeki, H. (1976) *J. Virol.* **18**, 1016–1023.
32. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
33. Yoshimori, R. N. (1971) Dissertation (Univ. California, San Francisco, CA).
34. Takanami, M. (1973) *FEBS Lett.* **34**, 318–322.
35. Richardson, C. C. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 158–165.
36. Helling, R. B., Goodman, H. M. & Boyer, H. W. (1974) *J. Virol.* **14**, 1235–1244.
37. McDonell, M. W., Simon, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110**, 119–146.
38. Yamawaki-Kataoka, Y., Sato, K., Shimizu, A., Kataoka, T., Mano, Y., Ono, M., Kawakami, M. & Honjo, T. (1979) *Biochemistry* **18**, 490–494.
39. Tilghman, S. M., Tiemeier, D. C., Seidman, J. G., Peterlin, B. M., Sullivan, M., Maizel, J. V. & Leder, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 725–729.
40. Brack, C. & Tonegawa, S. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5652–5656.
41. Francus, T. & Birshtein, B. K. (1978) *Biochemistry* **17**, 4324–4331.
42. Fougereau, M., Bourgois, A., Preval, C., Rocca-Serra, J. & Schiff, C. (1976) *Ann. Immunol. (Paris)* **127C**, 607–631.
43. Milstein, C., Adetugbo, K., Cowan, N. J. & Secher, D. S. (1974) in *Progress in Immunology*, eds. Brent, L. & Holborow, J. (North-Holland, Amsterdam), Vol. 1, pp. 157–168.
44. Birshtein, B. K., Hussain, Q. Z. & Cebra, J. J. (1971) *Biochemistry* **10**, 18–25.
45. Turner, K. J. & Cebra, J. J. (1971) *Biochemistry* **10**, 9–17.
46. Cebra, J. J., Steiner, L. A. & Porter, R. R. (1968) *Biochem. J.* **107**, 79–88.
47. Fruchter, R. G., Jackson, S. A., Mole, L. E. & Porter, R. R. (1970) *Biochem. J.* **116**, 249–259.
48. Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. P., Rutishauser, V. & Waxdal, M. J. (1969) *Proc. Natl. Acad. Sci. USA* **63**, 78–85.
49. Frangione, B. & Milstein, C. (1969) *Nature (London)* **224**, 597–599.
50. Franklin, E. C. & Frangione, B. (1971) *Proc. Natl. Acad. Sci. U. A* **68**, 187–191.
51. Cooper, S. M., Franklin, E. C. & Frangione, B. (1972) *Science* **176**, 187–189.
52. Frangione, B., Lee, L., Haber, E. & Bloch, K. J. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 1073–1077.
53. Robbinson, E. A., Smith, D. F. & Appella, E. (1974) *J. Biol. Chem.* **249**, 6605–6610.
54. Kunkel, H. G., Natvig, J. B. & Joslin, F. G. (1969) *Proc. Natl. Acad. Sci. USA* **62**, 144–149.
55. Natvig, J. B. & Kunkel, H. G. (1974) *J. Immunol.* **112**, 1277–1284.
56. Gilbert, W. (1978) *Nature (London)* **271**, 501.
57. Darnell, J. E., Jr. (1978) *Science* **202**, 1257–1260.
58. Sakano, H., Rogers, J. H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. & Tonegawa, S. (1979) *Nature (London)* **277**, 627–633.
59. Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 857–861.

Proc. Natl. Acad. Sci. USA
Vol. 77, No. 2, pp. 919–923, February 1980
Biochemistry

Rearrangement of immunoglobulin γ1-chain gene and mechanism for heavy-chain class switch

(cloning/nucleotide sequence/recombination site/S region/μ-chain gene)

TOHRU KATAOKA, TOSHIAKI KAWAKAMI, NAOKI TAKAHASHI, AND TASUKU HONJO*

Department of Physiological Chemistry and Nutrition, Faculty of Medicine, University of Tokyo, Tokyo 113, Japan

Communicated by Osamu Hayaishi, December 4, 1979

ABSTRACT Cloning and nucleotide sequence determination suggest that the rearranged γ1-chain gene in a γ1-chain-producing myeloma appears to be formed by the recombination between the 5′ flanking regions of the γ1- and μ-chain genes of undifferentiated cells. The recombination site is distinct from the putative J region and is a novel region that we call the S region. We have extended our previous model that explains the heavy-chain class switching by two or more successive recombination events.

A complete immunoglobulin gene is formed by a recombination event that brings the variable (V) gene into close proximity with the constant (C)-region gene (1–3). Studies (4–6) on cloned immunoglobulin light (L)-chain genes have unequivocally demonstrated that such recombination takes place between a germline V gene and a J-region gene that encodes 13 amino acids at the carboxyl end of the V region. In addition to the similar V–J recombination, immunoglobulin heavy (H)-chain genes must have unique features (absent from L-chain genes) that explain the phenomenon called H-chain class switch by which a given lymphocyte appears to be able to sequentially associate a single V region with two or more different classes of H-chain C regions.

We have presented evidence that deletion of specific C_H genes accompanies expression of the C_H gene in mouse myeloma cells (7). We proposed the allelic deletion model in which recombination of V_H and C_H genes occurs on one allele by deletion of the DNA segment that is located between the recombining V_H and C_H genes. Further, the C_H genes are proposed to be aligned in the order of V_H, spacer, μ, γ3, γ1, γ2b, γ2a, and α. We explained the H-chain class switch by successive allelic deletion events.

In this report we present evidence that the γ1-chain gene on one allele is rearranged in γ1-chain-producing myeloma cells whereas the γ1 gene on the other allele remains indistinguishable from the embryonic form. Comparison of the nucleotide sequences of the rearranged and embryonic γ1-chain genes indicates that the 5′ portion of the two genes are different but the other portions appear to be identical. The recombination site of the rearrangement is not the J region, as shown in L-chain genes, but a novel region referred to as the S region, which may be responsible for heavy-chain class switch. The sequence homologous to the 5′ new segment of the rearranged γ1-chain gene is found in the 5′ flanking region of the μ-chain gene in DNA from newborn mice.

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.

EXPERIMENTAL PROCEDURES

High molecular weight DNA was prepared from newborn mice or myeloma tumor and digested with *Eco*RI to completion. The fragments containing immunoglobulin genes were partially purified by RPC5 column chromatography (8) and agarose gel electrophoresis (9) as described (10, 11). The *Eco*RI fragments obtained were cloned with λgtWES (12) as an EK2 vector in a P2 facility. The recombinant DNA was packaged *in vitro* into phage coats (13), and the resultant plaques were screened by *in situ* hybridization (14) with the nick-translated DNA as probe. Recombinant phage DNA was prepared as described (15), and the insert was recloned into the *Eco*RI site of pBR322 (16).

Nucleotide sequence was determined by the method of Maxam and Gilbert (17), with slight modifications (11). Restriction DNA fragments were electrophoresed in 0.7% agarose gels (type I, Sigma) and transferred to nitrocellulose filters (18). The filter was hybridized to an appropriate probe as described (19) with slight modifications (10). R loop was examined by electron microscopy as described (10). Sources of restriction enzymes have been described (10, 11). [^{32}P]cDNA complementary to μ-chain mRNA was synthesized with oligo(dT) as primer (20). The mRNA used as template was partially purified by successive oligo(dT)-cellulose chromatography from TEPC 183 myeloma.

RESULTS AND DISCUSSION

γ1-Chain Gene Is Rearranged in γ1-Chain-Producing Myelomas. Total cellular DNAs extracted from various myelomas and newborn mice were digested with *Eco*RI, electrophoresed in an agarose gel, and transferred to nitrocellulose filters. By using the nick-translated γ1-chain gene fragment (IgH2) cloned from *Eco*RI digests of newborn mouse DNA (11) as a hybridization probe, we detected DNA restriction fragments containing the γ1-chain gene in various DNAs. As shown in Fig. 1, a 6.6-kilobase (kb) fragment equivalent to IgH2 was found in all the DNA preparations tested. The γ1-chain-producing myelomas, MC101 and MOPC 31C, however, had extra bands of 8.3 and 4.0 kb, respectively. The 8.3-kb band of MC101 DNA was also observed with the nick-translated γ1-chain cDNA clone pG1-6 (22) as probe. The 6.6-kb bands are faint in DNA of the γ2b-chain-producing myelomas, indicating that the γ1-chain gene was deleted in these myelomas

Abbreviations: C and V regions, constant and variable regions, respectively; H and L chains, heavy and light chains, respectively; kb, kilobase.

* Present address: Department of Genetics, Osaka University School of Medicine, Osaka 530, Japan.

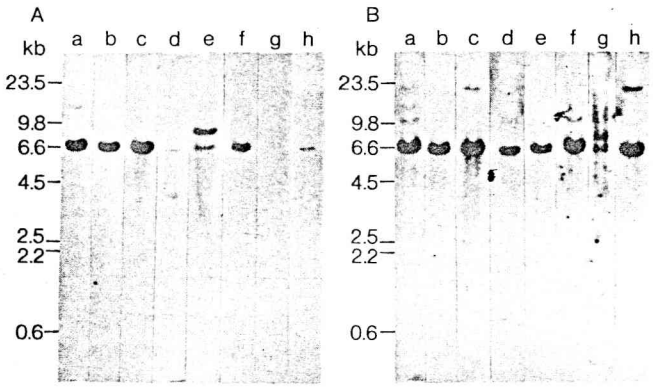


FIG. 1. Identification of *Eco*RI fragments of myeloma DNAs that contain the γ1- or γ2b-chain gene. *Eco*RI-digested DNAs (10 μg each) of myelomas and newborn mice were electrophoresed in 0.5% agarose gels and transferred to nitrocellulose filters (Schleicher & Schuell). The filters were hybridized with [^{32}P]IgH2 or [^{32}P]IgH22 (specific activity, 2×10^8 cpm/μg) labeled by nick-translation (21), washed, and autoradiographed. Shown are autoradiograms of the Southern blots hybridized with [^{32}P]IgH2 (*A*) and [^{32}P]IgH22 (*B*). *Eco*RI-digested DNAs of myelomas (H-chain-producing): lanes a, newborn mice; lanes b, MOPC 104E (μ); lanes c, J606 (γ3); lanes d, MOPC 31C (γ1); lanes e, MC 101 (γ1); lanes f, MOPC 70A (γ1); lanes g, MOPC 141 (γ2b); lanes h, MPC 11 (γ2b).

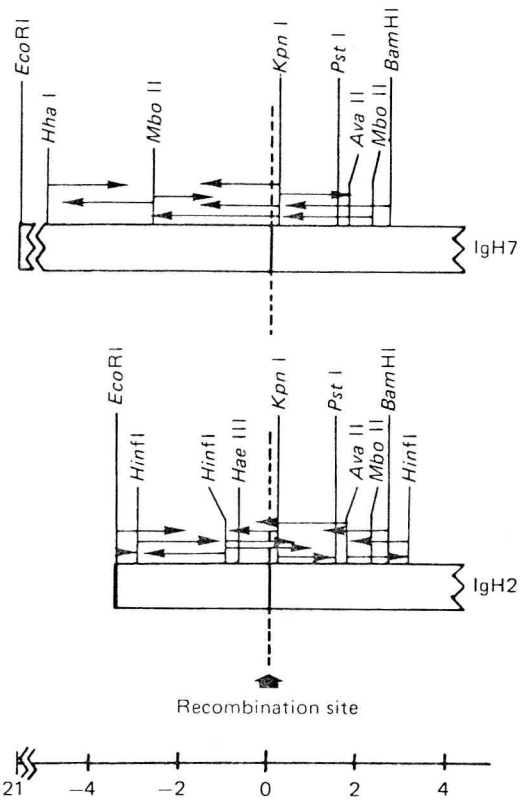


FIG. 2. Comparison of restriction enzyme cleavage maps of IgH7 and IgH2 and strategy for nucleotide sequence determination. Detailed restriction maps of the 5′-terminal 600 bases of IgH2 and the corresponding region of IgH7 are shown. Sites for phosphorylation and range and direction of sequences read are shown by horizontal arrows. DNAs of the hybrid plasmids were isolated as described (23). Restriction cleavage sites were mapped by combined cleavage of the plasmids with various restriction enzymes or by partial digestion of DNA fragments labeled at one of their 5′ termini with T4 polynucleotide kinase and [γ-^{32}P]ATP (24).

in agreement with our previous results (7). When the nick-translated γ2b-chain gene clone IgH22 (10) was used as probe, DNAs of all myelomas except MOPC 141 showed a single 6.6-kb fragment equivalent to IgH22. DNA of the γ2b-chain-producing myeloma MOPC 141 showed an extra band of 7.8 kb in addition to the 6.6-kb band. Although *Eco*RI digestion of MPC 11 DNA produced the 6.6-kb band of the γ2b-chain gene, we have recently cloned a partially digested *Eco*RI fragment (13 kb) from MPC 11 DNA and shown that there is a rearrangement upstream from the 6.6-kb fragment of one allele (unpublished data). We suspect a similar rearrangement in the γ1-chain gene of MOPC 70A that produced the 6.6-kb *Eco*RI fragment of the γ1-chain gene. These results indicate that the γ1 (γ2b)-chain gene is specifically rearranged in the γ1 (γ2b)-chain-producing myelomas and that the site of rearrangement varies between myelomas.

The 6.6-kb γ1-chain gene of MC 101 DNA was indistinguishable from IgH2 by digestion with *Bam*HI, *Sac* I, *Hha* I, *Xba* I, *Kpn* I, or *Hin*dIII (data not shown). We refer to the 6.6-kb γ1-chain gene fragment (equivalent to IgH2) in myelomas and newborn mice as the embryonic gene and the 8.3- and 4.0-kb γ1-chain gene fragments in myelomas as the rearranged genes. These results indicate that the H-chain genes expressed in myelomas have undergone a rearrangement that takes place on only one of the alleles although we cannot exclude the possibility that the intact allele has a different rearrangement outside the 6.6-kb fragment which we did not detect.

Cloning and Characterization of Rearranged γ1-Chain Gene. The rearranged γ1-chain gene of MC101 was cloned by using nick-translated pG1-6 as a probe. Three independent positive clones were obtained after about 1.2×10^5 plaques were screened. DNAs of the three clones were extracted and shown to hybridize with γ1-chain [^{3}H]cDNA in solution. One clone was designated λgt*WES*·IgH7 (abbreviated IgH7), and the insert DNA was recloned into a plasmid pBR322, which was called pIgH7.

Then, IgH7 DNA was characterized by restriction enzyme cleavage and R-loop formation. The restriction maps of the IgH7 and IgH2 inserts were compared (Figs. 2 and 5). The 3′-terminal, 6.3-kb portions of the IgH7 (8.3-kb) and IgH2 (6.6-kb) inserts were indistinguishable from each other by restriction cleavage sites. The 5′-terminal portions of the IgH7 and IgH2 inserts were different from each other with respect to locations of various restriction sites.

R loops formed between IgH7 DNA and MC101 γ1-chain mRNA had shapes identical to those formed between IgH2 DNA and MOPC 31C γ1-chain mRNA, showing intervening sequences between domains (10, 11, 25). No R loops corresponding to the V gene of MC101 γ1 chain was observed. These results indicate that a new DNA segment (about 2 kb) replaced the 5′-terminal portion (about 300 base pairs) of the embryonic γ1-chain gene to yield the rearranged γ1-chain gene of 8.3 kb.

Comparison of Nucleotide Sequences of Rearranged and Embryonic γ1-Chain Genes. Detailed restriction enzyme cleavage maps were constructed and the nucleotide sequences were determined to characterize the new DNA segment introduced to IgH7. Fig. 2 shows the detailed restriction maps and the sequencing strategy of the 5′-terminal portions to the *Bam*HI sites, which are located 6.0 kb from the 3′ ends of the IgH7 and IgH2 inserts. The two clones share common restriction enzyme cleavage sites between the *Bam*HI sites and the *Kpn* I sites, which lie 2.0 kb and 0.35 kb from the 5′ ends of the IgH7 and IgH2 inserts, respectively. However, a *Hae* III site 0.28 kb away from the 5′ end of IgH2 is absent from IgH7 at the corresponding position relative to the *Kpn* I site. Other cleavage sites upstream from the *Hae* III site are different in

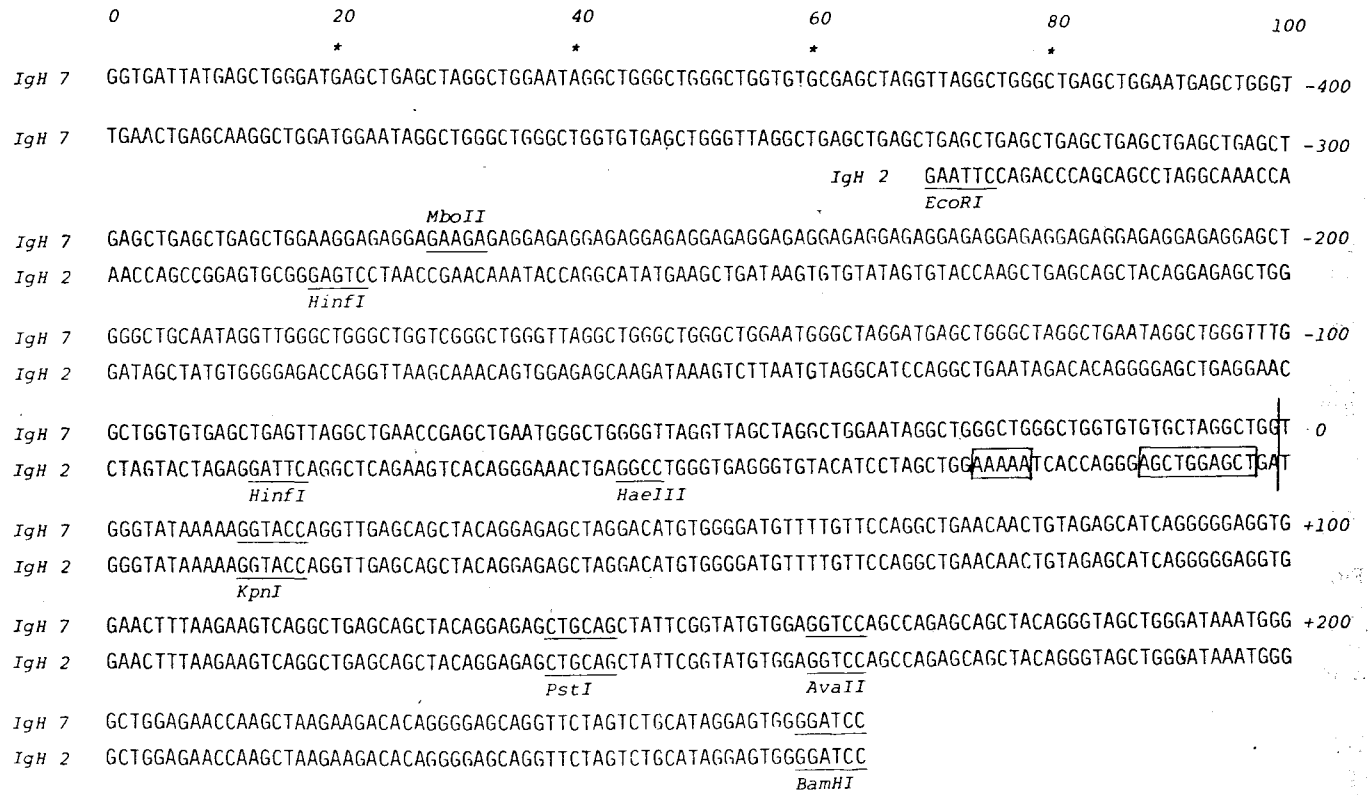


FIG. 3. Nucleotide sequences of the 5′-terminal portions of IgH7 and IgH2. The nucleotide sequences of the 5′-terminal 600-base-pair segment of IgH2 and the corresponding region of IgH7 are displayed with the direction of the transcription of the structural gene. A vertical line shows the site of recombination. A palindrome sequence A-G-C-T-(G-G)-A-G-C-T and a pentanucleotide A-A-A-A-A, which are adjacent to the recombination site, are boxed.

the two clones. Comparison of the nucleotide sequences of IgH7 and IgH2 (Fig. 3) unambiguously locates the recombination site of the rearranged gene at base 17 before the *Kpn* I site. The nucleotide sequences of the region (≈270 bases) from the recombination site to the *Bam*HI site of IgH7 and IgH2 match each other completely. The two sequences upstream from the recombination site, however, diverge from each other. Tandem repetitive sequences of $(G\text{-}A\text{-}G\text{-}C\text{-}T)_n$, $(G\text{-}A\text{-}G\text{-}A\text{-}G)_n$, and $(G\text{-}G\text{-}C\text{-}T\text{-}G)_n$ are prominent in the ≈400-base-pair-long region before the recombination site in IgH7. No such sequences were found in the corresponding region of IgH2.

We searched for the base sequences corresponding to the V region of MC101 γ1 chain. Though the amino acid sequence of the V region of MC101 γ1 chain was not determined, the homology in the framework regions shared among many V-region sequences of other known H chains (26–30) enabled us to identify a possible V-gene sequence. Six COOH-terminal amino acids (Try-Gly-X-Gly-Thr-X-X-Thr-Val) are especially well conserved (26–30) and have been assigned as a part of the putative J region in H-chain proteins (30). We have determined the sequence of almost the entire 5′-terminal portion of IgH7 and failed to find these conserved sequences, confirming an earlier observation that no R loops corresponding to the V gene were formed. The results indicate that rearrangement of the γ1-chain gene requires a recombination other than a direct recombination between a V gene and a J region that is a single genetic event to complete the λ- or κ-type L-chain genes (2–6).

5′-Terminal Portion of Rearranged γ1-Chain Gene Is Linked to μ-Chain Gene in DNA of Newborn Mice. To search for the origin of the 5′-terminal segment of IgH7 in DNA of undifferentiated cells, we tested *Eco*RI-digested DNA of newborn mice by *in situ* hybridization by using nick-translated *Bam*HI fragment B of pIgH7 (*Bam* B fragment), which contained the 5′-terminal 2.2-kb DNA segment of IgH7 and 0.38-kb DNA segment derived from pBR322, as a probe. The *Bam* B fragment has a sequence about 200 base pairs long common to IgH2. DNA from newborn mice contained three bands 13, 10, and 6.6 kb long which hybridized with the *Bam* B fragment of pIgH7. The three fragments were purified by preparative agarose gel electrophoresis and cloned by using λgtWES. The clones that contained the inserts of 13, 10, and 6.6 kb were designated IgH701, IgH703, and IgH705, respectively. IgH705 was shown to be identical to IgH2 by restriction mapping and *in situ* hybridization.

Locations of the sequence homologous to the *Bam* B fragment were determined by *in situ* hybridization of restriction DNA fragments (Fig. 4*A*). When digested with *Eco*RI, *Hin*dIII, and *Bam*HI/*Eco*RI, DNA of IgH701 produced one band (13 kb), three bands (24, 3.8, and 0.7 kb), and one band (7.2 kb), respectively, that hybridized with the *Bam* B fragment. Similarly, digestion of IgH703 with *Eco*RI, *Hin*dIII, and *Xba* I produced 10-, 10.3-, and 3.4-kb fragments, respectively, hybridizing with the *Bam* B fragment of IgH7. As summarized in Fig. 5, the homologous sequences were located at the end of IgH701 close to the *Eco*RI site and at the middle of IgH703. Comparison of the restriction cleavage maps of IgH7, IgH701, and IgH703 indicated that the 5′-end portion (about 1.8 kb) of IgH701 is identical to the 5′ end of IgH7. The relative locations of *Eco*RI, *Xba* I, *Sac* I, *Hin*dIII, and *Hha* I cleavage sites in the 5′-terminal 1.8-kb portion were conserved between IgH7 and IgH701, except for a *Sac* I site located 1.8 kb away from the 5′ end of IgH701. The *Kpn* I and *Bam*HI cleavage sites of IgH7, which were shared by IgH2, were absent from IgH701 (Fig. 5). The portion of IgH703 that hybridized to the *Bam* B fragment was significantly different from the *Bam* B fragment of

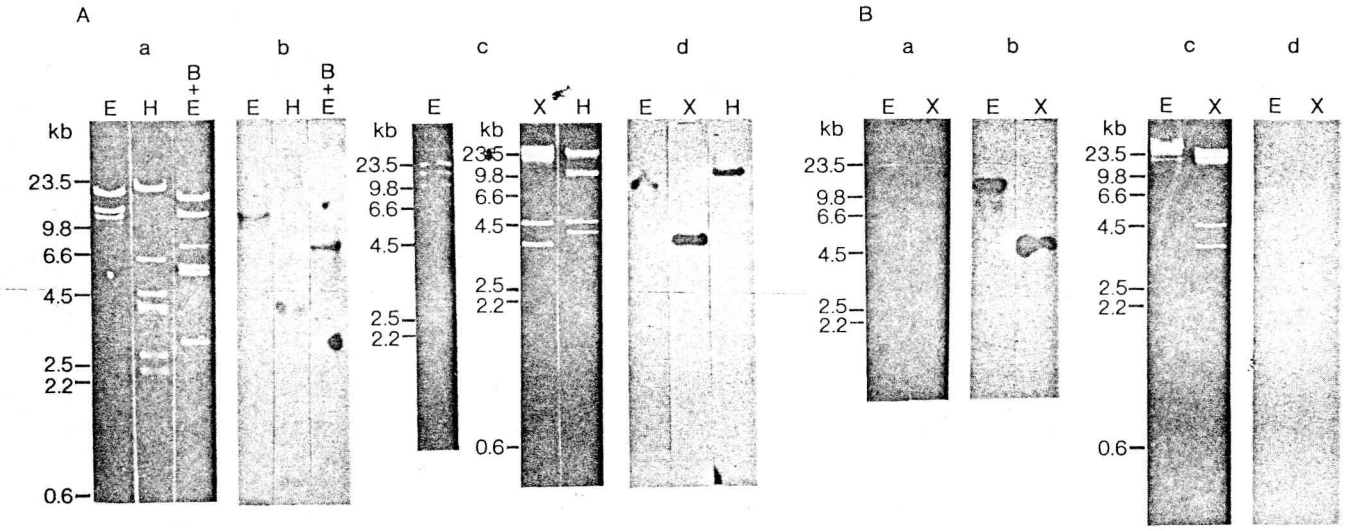


FIG. 4. *In situ* hybridization of IgH701 and IgH703 restriction DNA fragments with the ^{32}P-labeled *Bam* B fragment of pIgH7 and μ-chain [^{32}P]cDNA. IgH701 and IgH703 DNAs (0.3 μg of each) were cleaved with the restriction enzymes indicated (E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; X, *Xba* I) and electrophoresed in 0.7% agarose gels. The Southern blot of the gel was hybridized with ^{32}P-labeled *Bam* B fragment (*A*) or μ-chain [^{32}P]cDNA (*B*). Ethidium bromide stains and autoradiogram of Southern blots of gels are shown: lanes a and b, IgH701; lanes c and d, IgH703.

IgH7. It is possible that only a small portion of the *Bam* B fragment sequence is shared by IgH703.

We then tested the possibility that other H-chain genes might be linked to the embryonic clones IgH701 and IgH703. DNAs of these clones were digested with *Eco*RI and *Xba* I, electrophoresed in 0.7% agarose gels, and hybridized with ^{32}P-labeled H-chain cDNA of various classes after Southern blotting of the gel. As shown in Fig. 4*B*, the μ-chain [^{32}P]cDNA hybridized specifically with certain restriction fragments of IgH701. No hybridization was observed with IgH703. Combination of the restriction map of IgH701 and *in situ* hybridization with the μ-chain cDNA indicated that the μ-chain gene lies between approximately 6 and 8 kb from the 5′ end of IgH701 (Fig. 5). We isolated the 1.2-kb *Hin*dIII fragment of IgH701, digested it with *Hin*fI, and determined the nucleotide sequences of more than 300 bases that match the amino acid sequences of the μ chain (31). For example, we found the nucleotide sequence

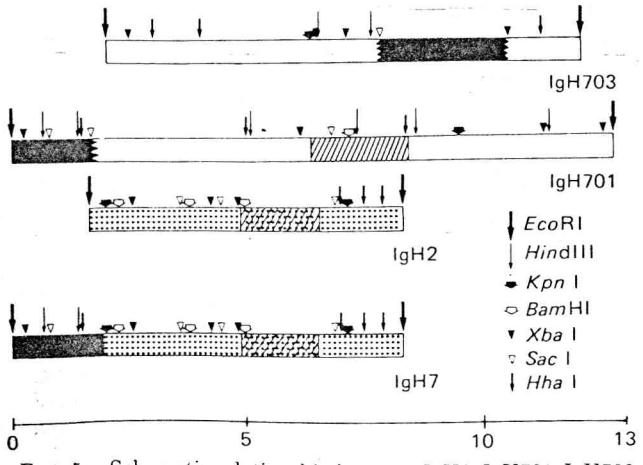


FIG. 5. Schematic relationship between IgH2, IgH701, IgH703, and IgH7. Restriction cleavage maps of IgH2, IgH701, IgH703, and IgH7 inserts are compared. The restriction map was constructed as described in the legend to Fig. 2. Dotted and filled segments indicate IgH2 (and its equivalent) and the newly introduced segment of IgH7 (and the region hybridizing with it), respectively. Oblique regions show the μ, γ1, and γ1 chain genes in IgH701, IgH2, and IgH7, respectively.

of CCA-CAG-AAG-AAA-TTC-ATC-TCA-AAA-CCC-AAT-GGT-AGG-TAT-CCC-CCC, only the first 30 bases of which coincide with the μ-chain amino acid sequence (31) of Pro-Gln-Lys-Lys-Phe-Ile-Ser-Lys-Pro-Asn (residues 436–445) corresponding to the end of the CH3 domain. The nucleotide sequence not only demonstrates unambiguously that IgH701 contains the μ-chain gene, but also indicates that the μ-chain gene is interrupted by an intervening sequence at the junction of the CH3 and CH4 domains as shown in other H-chain genes (10, 11, 25, 32). These results led us to conclude that the DNA segment proximal to the 5′ end of the μ-chain gene in DNA from newborn mice has recombined with the sequences flanking the 5′ end of the γ1-chain gene in the γ1-chain-producing myeloma MC101.

S Region. In view of the fact that such rearrangement takes place specifically in the γ1-chain gene in the γ1-chain producing myeloma, it is likely that this rearrangement is responsible for the class-switch expression of the γ1-chain gene. We postulate that IgH7 is the 3′ part of the completed γ1-chain gene that is linked to the V gene and expressed in MC 101 myeloma although we do not find the V-region sequence in IgH7. Early *et al.* (32) found that the α-chain gene from MOPC 603 has a 6.8-kb intervening sequence between the V- and C-region sequences; thus it is not surprising that the V-region sequence is not present in IgH7. We cloned the 7.8-kb rearranged γ2b-chain gene of MOPC 141 myeloma (Fig. 1), the 5′ portion of which seems to be identical to the 5′ portion of IgH701. The results suggest that such rearrangement is universal and not an abnormal rearrangement specific to MC 101 myeloma. We define the DNA segment responsible for the class-switch recombination as switch (S) region. We assume that the DNA segments adjacent to the recombination sites of IgH2 and IgH701 contain the S regions. The S regions originally linked to the μ- and γ-chain genes are referred to as the Sμ and Sγ regions, respectively. We presume that the class-switch recombination takes place by a looping-out mechanism forming a stem structure between the Sμ and Sγ sequences in a fashion analogous to the V_L–J_L recombination in which the two identical palindrome sequences C-A-C-(A)-G-T-G adjacent to the recombination sites of the V_K- and J_K-region sequences are assumed to form an inverted repeat-stem structure to loop out

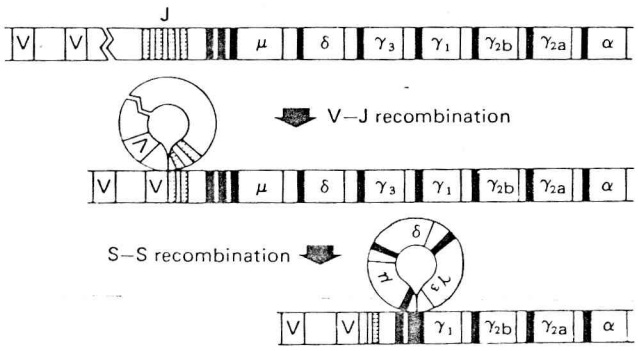


FIG. 6. A model for heavy-chain class switch. Recombination events producing a complete γ1-chain gene are schematically represented. The alignment of V_H and C_H genes is based on our previous model (7), in which the δ gene was omitted. Dotted segments show the J regions. Filled segments before the C_H genes (μ, γ1, α, etc.) indicate the S region of the C_H genes (Sμ, Sγ1, Sα, etc.). Both V–J and S–S recombinations are accompanied by deletion of in-between DNA segments as proposed (7).

the intermediate DNA segment (4, 5). We found a palindrome sequence A-G-C-T-(GG)-A-G-C-T at the portion immediately adjacent to the recombination site of the embryonic γ1-chain gene (IgH2). There is a pentanucleotide, A-A-A-A-A, 22 base pairs away from the recombination site of IgH2. The pentameric A·T base pairs were found in the inverted repeat-stem structures of the V_L–J_L recombination and the bacterial insertion element IS1 (4, 5, 33).

Model for Heavy-Chain Class Switching. During differentiation of a B lymphocyte, a single V_H gene is first expressed as the constituent of the μ chain; later the C part of the expressed H chain switches to γ or α with the V part unchanged. We extend our previous model (7) based on the recent findings. The V_L gene recombines with the J-region segment located close to the κ- or λ-chain C gene (2–6). It is reasonable to assume a similar structural basis for joining the V_H and C_H genes because recent amino acid sequence data of V_H regions have presented evidence for the existence of the J segment corresponding to the 3′-terminal portion of the variable region (29, 30). If this is so, as the first step of H-chain gene expression, a selected V gene recombines with one of the J regions that may be clustered at the region flanking the 5′ end of the Sμ region, resulting in expression of the μ-chain gene (Fig. 6). We have demonstrated that the γ1-chain gene undergoes recombination at the S region, which is distinct from the J region. As predicted by our previous model (7), the class switch from the μ-chain gene to the γ-chain gene requires a second recombination event which takes place between the Sμ and Sγ regions (Fig. 6).

There may be several Sμ regions between the J and μ-chain gene sequences and other sets of S regions at the 5′ end of each heavy-chain C gene (Sγ, Sα, Sδ, and Sε). Each Sμ region may be designed to recombine with the S regions of the specific C_H gene. It is also possible that each J is followed by an Sμ region(s). We postulate multiple S regions before each C_H gene because *Eco*RI fragments of the rearranged γ1- and γ2b-chain genes in different myelomas are of varied lengths (Fig. 1). We have obtained preliminary evidence that IgH703, but not IgH701, contains a sequence homologous to the small region (a few hundred base pairs) immediately before the recombination site of IgH7. The results suggest that a third recombination may be necessary to complete the rearranged γ1-chain gene. This model is essentially an elaborated version of our previous model (7) and can easily explain the phenomenon that the V region remains unchanged after the class switch.

We thank Dr. Y. Mano for his continuous encouragement. We also thank Dr. M. Potter for providing myelomas used in this report, except for MC 101, which was kindly supplied by Dr. S. Migita. This investigation was supported in part by grants from the Ministry of Education, Science, and Culture of Japan.

1. Drayer, W. J. & Bennet, J. D. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 864–869.
2. Brack, C., Hirama, M., Lenhard-Schuller, R. & Tonegawa, S. (1978) *Cell* **15**, 1–14.
3. Seidman, J. G. & Leder, P. (1978) *Nature (London)* **276**, 790–795.
4. Max, E. E., Seidman, J. G. & Leder, P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3450–3454.
5. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. (1979) *Nature (London)* **280**, 288–294.
6. Seidman, J. G., Max, E. E. & Leder, P. (1979) *Nature (London)* **280**, 370–375.
7. Honjo, T. & Kataoka, T. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2140–2144.
8. Pearson, R. L., Weiss, J. F. & Kelmers, A. D. (1971) *Biochim. Biophys. Acta* **228**, 770–774.
9. Polsky, F., Edgell, M. H., Seidman, J. G. & Leder, P. (1978) *Anal. Biochem.* **87**, 397–410.
10. Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. & Honjo, T. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4240–4244.
11. Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N. & Mano, Y. (1979) *Cell* **18**, 559–568.
12. Leder, P., Tiemeier, D. & Enquist, L. (1977) *Science* **196**, 175–177.
13. Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Richard, J. E. Slighton, J. L., Tucker, P. W. & Smithies, O. (1978) *Science* **202**, 1279–1283.
14. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
15. Tiemeier, D. C., Tilghman, S. M. & Leder, P. (1977) *Gene* **2**, 173–191.
16. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heynecker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95–113.
17. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
18. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
19. Jeffrey, A. J. & Flavell, R. A. (1977) *Cell* **12**, 429–439.
20. Honjo, T., Packman, S., Swan, D. & Leder, P. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3659–3663.
21. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184–1188.
22. Obata, M., Yamawaki-Kataoka, Y., Takahashi, N., Kataoka, T., Shimizu, A., Mano, Y., Seidman, J., Peterlin, M., Leder, P. & Honjo, T. (1980) *Gene*, in press.
23. Clewell, D. & Helinski, D. (1969) *Proc. Natl. Acad. Sci. USA* **62**, 1159–1166.
24. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2398.
25. Sakano, H., Rogers, J. H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. & Tonegawa, S. (1979) *Nature (London)* **277**, 627–633.
26. Fougereau, M., Bourgois, A., Preval, C., Rocca-Serra, J. & Schiff, C. (1976) *Ann. Immunol. (Paris)* **127C**, 607–631.
27. Barstad, P., Farnsworth, V., Weigert, M., Cohn, M. & Hood, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4096–4100.
28. Adetsugbo, K. (1978) *J. Biol. Chem.* **253**, 6068–6075.
29. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2429–2433.
30. Rao, D. N., Rudikoff, S., Krutzsch, H. & Potter, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2890–2894.
31. Kehry, M., Sibley, C., Fuhrman, J., Schilling, J. & Hood, L. E. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2932–2936.
32. Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 857–861.
33. Ohtsubo, H. & Ohtsubo, E. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 615–619.