The EMBO Journal vol.10 no.3 pp.505-511, 1991

Evolution of immunoglobulin light chain genes: analysis of *Xenopus* IgL isotypes and their contribution to antibody diversity

Joseph Schwager, Niels Bürckert, Martine Schwager and Melanie Wilson

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

Communicated by H. von Boehmer

The amphibian *Xenopus laevis* expresses several types of immunoglobulin light chain (IgL). cDNA clones for two IgL isotypes, $C_\sigma1$ and $C_\sigma2$, were analysed. C_σ is expressed in spleen and mitogen-stimulated B cells, like another *Xenopus* IgL type, termed C_ϱ. C_σ shares <33% residues with C_ϱ or with C_L regions of shark, chicken and mammals. This suggests that C_σ diverged from a common ancestor of C_L regions before or at the emergence of amphibians. Two families of V_L elements, $V_\sigma1$ and $V_\sigma2$ are associated with C_σ (but not with C_ϱ). They rearrange to their own set of J_L elements, $J_\sigma1$ and $J_\sigma2$, which are poorly related to other J elements of the Ig gene family. The *Xenopus* genome contains a few $V_\sigma2$ and multiple $V_\sigma1$ elements (comparable with mammalian V_χ), but only two C_σ genes. Thus, the organization and expression of *Xenopus* IgL loci are apparently similar to mammalian IgL loci but different from shark and chicken IgL loci. Only a few V_L elements are expressed, since cDNA clones show extensive sharing of CDR1 and CDR2 sequences; some clones differ only in CDR3. Rearranging V_L and J_L elements increases CDR3 diversity in both V_σ families, but abortive rearrangements are frequent in $V_\sigma1$ regions. The very poor heterogeneity of expressed V_L elements therefore appears to limit antibody diversity in *Xenopus*.

Key words: antibody diversity/evolution/Ig light chains/*Xenopus*

Introduction

The basic unit of every immunoglobulin (Ig) molecule consists of two identical Ig heavy chains (IgH) and two identical light chains (IgL). In mammals, the organization and expression of multiple IgH and IgL genes is well documented (for a review see e.g. Alt *et al.*, 1987). The Ig loci of lower vertebrates are apparently less complex than those in mammals, as reflected by the detection of only one IgH class in shark, two IgH classes in bony fish and urodele amphibians and three IgH classes in anuran amphibians and birds (for a review see Du Pasquier, 1989). Similarly, the diversity of antibody populations tends to increase from lower to higher vertebrates (reviewed by Du Pasquier, 1982). The unique structure of the Ig loci in elasmobranchs (Hinds and Litman, 1986; Shamblott and Litman, 1989) or the low intrinsic complexity of the variable elements in amphibians (Schwager *et al.*, 1989) may partially account for these particular phenomena.

Mammalian Ig molecules contain IgL chains of the κ or λ type; their relative expression differs from one species to another. Extant species have IgL chains which reflect different evolutionary pathways. While chicken IgL chains are of λ-type (Reynaud *et al.*, 1985), those in elasmobranchs are either only remotely related to mammalian IgL chains (Shamblott and Litman, 1989) or apparently related to human λ chains (Schluter *et al.*, 1989). The anuran amphibian *Xenopus laevis* produces very heterogeneous IgL populations which are distinguishable by SDS-PAGE analysis (Hadji-Azimi, 1975; Schwager and Hadji-Azimi, 1984, 1985); their association with IgM, IgX and IgY is highly non-random (Hsu and Du Pasquier, 1984; Hsu *et al.*, 1985). One IgL chain of *Xenopus* has both κ- and λ-like features (Zezza *et al.*, 1988, and manuscript submitted). In shark, chicken and mammals, Ig light and heavy chains are diversified by the same distinct species-specific mechanisms (Tonegawa, 1983; Reynaud *et al.*, 1985, 1989; Shamblott and Litman, 1989) which suggests that both loci have been subjected to similar evolutionary constraints. It is therefore reasonable to expect that in *Xenopus* the IgL locus and its expression shares features with those described for IgH chains (Schwager *et al.*, 1989).

In order to gain insight into the evolutionary diversity of the IgL chains and the particular features of the Ig expression in *Xenopus* (for a review see Du Pasquier *et al.*, 1989), we have isolated different *Xenopus* IgL cDNA clones. We present here the sequence of a novel IgL type and its association with V_L elements. Unlike what might have been expected in view of the gel analysis of IgL chains, the expressed V_L elements are not very diverse. These data provide more information about IgL evolution and the limited diversity of antibodies in *Xenopus*.

Results

Isolation of cDNA clones from different Xenopus IgL chains

Rabbit anti-*Xenopus* IgM and anti-*Xenopus* IgY antibodies react not only with their respective heavy chain class but also with the light chain type which is preferentially associated with that class. Such antisera were used in a differential immunoscreening procedure to isolate *Xenopus* IgL chains from an amplified cDNA library, prepared from mitogen-stimulated spleen cells. About 100 clones reacting with anti-IgM were identified and subsequently hybridized with a C_μ probe (Schwager *et al.*, 1988a). An anti-IgM^+, C_μ^- clone, i.e. a putative IgL cDNA clone contained a 176 nucleotide (nt) insert that encodes part of an Ig constant region domain. The cDNA library was rescreened with this probe, and numerous other clones many of which had inserts of >1 kb were isolated. (We will refer to the sequence of these cDNA clones encoding the variable and constant regions as V_σ and C_σ, respectively.) A

© Oxford University Press

subsequent round of screening was made with inserts of cDNA clones that contain $V_\sigma1-C_\sigma1$ or $V_\sigma2-C_\sigma1$ (see below). Similarly, a series of clones were isolated by immunoscreening cDNA libraries with anti-IgY (Zezza *et al.*, 1988 and manuscript submitted). We refer to the protein sequence of the constant region deduced from these cDNA clones as C_ϱ.

Expression of Xenopus IgL classes

We analysed the tissue expression of C_σ in comparison with C_ϱ on Northern blots. Both *Xenopus* IgL types are expressed strongly in spleen, weakly in liver or thymus and not at all in ovary (Figure 1). Low-density spleen cells, which synthesize IgM and IgY, and high-density spleen cells, which mainly synthesize membrane IgM (Schwager, 1983) produce ϱ mRNA (Figure 1A, lanes 7 and 8) and σ mRNA (data not shown). Spleen cells stimulated with pokeweed mitogen (PWM) or lipopolysaccharide (LPS), which produce IgM and IgX (Schwager and Hadji-Azimi, 1984; Hsu *et al.*, 1985), also express both IgL types. The cDNA probes for the two IgL types hybridize with RNA species of ~1 kb in size from different lymphoid tissues or mitogen-stimulated cells.

Comparison of Xenopus C_σ with other C_L regions

We have partially or completely sequenced the C_L region of more than 20 C_σ^+ cDNA clones. They form two groups, $C_\sigma1$ and $C_\sigma2$, which differ in 26 nucleotides in the coding region (Figure 2). This results in nine substitutions at the amino acid level. The untranslated 3′ sequence of $C_\sigma1$ is 318 nt in length with a polyadenylation signal sequence (AATAAA) at 301 nt 3′ to the termination codon. The untranslated 3′ end of $C_\sigma2$ is 358 nt in length, with an AATAAA sequence 17 nt before the poly(A) tail. Since there is only 80% sequence identity between $C_\sigma1$ and $C_\sigma2$ in their untranslated 3′ end, they only weakly cross-hybridize under stringent hybridization conditions. Most $C_\sigma1$ cDNA clones were associated with $V_\sigma1$ or $V_\sigma2$ regions (see below), while the two $C_\sigma2$ cDNA clones did not contain V_σ regions.

C_σ shares only 29% residues with C_ϱ and it has 29–33% protein sequence identity with the C_L regions of shark, chicken or mammals (Figure 3). In almost all species, a few C_L residues are invariant: Pro119, Pro120, Cys134, Trp148, Asp151, Thr162, Ser176, Ser177, Leu179, Cys194, Cys214. Most of these and other common residues are located in the antiparallel β-pleated sheet of the Ig fold (Figure 3). *Xenopus* C_σ, C_ϱ and the C_L of other vertebrates (see Figure 3) are most conserved in the four-stranded β-pleated sheet and, in particular, in and around 4–3. This β-pleated sheet forms the contact area between C_L and C_H (for a review see Hasemann and Capra, 1989). All nine amino acid differences between $C_\sigma1$ and $C_\sigma2$ are located outside the four-stranded β-pleated sheet (not shown); therefore, they should not affect the C_L-C_H interaction.

Particular features of variable regions associated with Xenopus C_σ

All σ^+ cDNA clones containing the complete or part of the variable (V_L) region could be grouped into two V_L families, $V_\sigma1$ and $V_\sigma2$. They are only found in association with C_σ. C_ϱ is associated with a different specific set of V_L regions (our unpublished observations). $V_\sigma1$ and $V_\sigma2$ have the landmarks of the V_L of other vertebrates, including the two

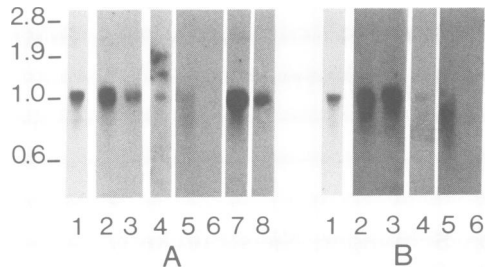

Fig. 1. Expression of two IgL classes in *Xenopus* lymphoid cells. Northern blot analysis of 10–20 μg total RNA extracted from total spleen (lane 1), from spleen cells stimulated for 6 days with PWM (lane 2) or with LPS (lane 3), from thymus (lane 4), liver (lane 5), ovary (lane 6), low density (lane 7), and high-density spleen cells (lane 8). Hybridization was done with a probe containing $C_\sigma1$ constant region, i.e. clone $\sigma6$ (see Figure 2) (panel B) or with a cDNA clone ($\varrho136$) containing the complete C_ϱ (panel A). RNA markers (in kb) are indicated at the left.

cysteine residues [Cys23 and Cys88 (Kabat *et al.*, 1987)] that form the intradomain S–S bond (Figure 4A and B). The conserved residues are located in the β-pleated sheet, in the turns or among the residues that participate in V_L-V_H interactions (Table I, Figure 4C).

However, V_σ regions have unusual features compared with V_L regions of other vertebrates. The prototype $V_\sigma1$ and $V_\sigma2$ (Figure 4) contain 115 residues rather than the ~110 residues of V_L regions in other vertebrates. The two cysteine residues forming the intradomain S–S bond are 73 residues apart, rather than the usual 65 residues. The otherwise invariant Trp36 and Trp47 (Kabat *et al.*, 1987) are substituted by Ile or Leu. Based on the conserved residues in the Ig fold (Table I), the CDR1 has a size similar to those of other V_L regions, but the region of the putative CDR2 loop (i.e. usually 6–7 residues) is longer in the V_σ regions (see Table I and Figure 4). FR4 of V_σ regions have only three to four residues in common with murine FR4, encoded by J_κ or J_λ elements, and they share only the Phe98 with the FR4 sequence of V_L regions associated with C_ϱ (Figure 5). Gly99 and Gly101, which are strictly conserved in J elements of IgH, IgL and T cell receptor (TCR) genes (Kabat *et al.*, 1987), are replaced by serine (or arginine) residues in all variable regions associated with C_σ. Two groups of FR4 sequences in V_σ regions differ by two residues at position 100 and 101.

Extensive sharing of CDR in V_σ regions

V_σ regions obtained from outbred animals share CDR sequences far more extensively than do V_H elements (Schwager *et al.*, 1989). Among the 12 $V_\sigma1$ regions shown in Figure 4C, three ($\sigma108$, $\sigma3$, $\sigma105$) are identical in CDR1 and CDR2; they differ only in CDR3. They share the same leader peptide and they may be transcripts from the same V_σ element, or a set of closely related $V_\sigma1$ elements. Clones $\sigma103$, $\sigma111$ and $\sigma14$ are transcripts from another V_σ element. Similarly, three other $V_\sigma1$ regions (clones $\sigma21$, $\sigma40$, $\sigma8$) represent a third expressed member of the $V_\sigma1$ family or a set of almost identical $V_\sigma1$ elements. Clone $\sigma28$ differs in CDR1 and CDR2 from the other $V_\sigma1$ regions. Clone $\sigma102$ differs from the prototype $V_\sigma1$ regions by 14 residues in the FRs.

The expressed $V_\sigma2$ elements are virtually identical in the

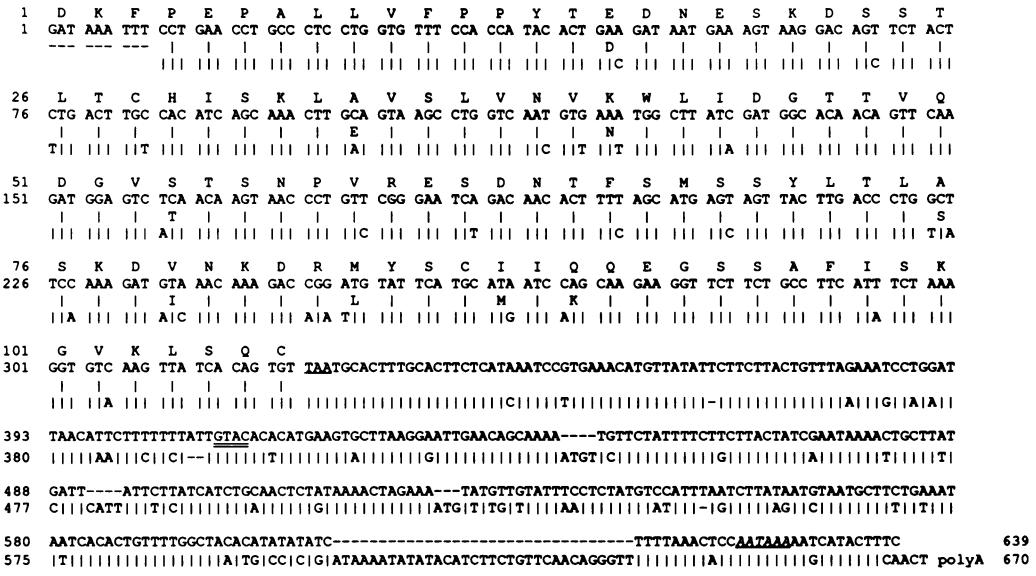

Fig. 2. Nucleotide and deduced protein sequence of the constant region of *Xenopus* Ig σ chains. The nucleotide and protein sequence (in the one-letter code) of clone σ6 (which starts in FR4) are given in the upper two lines. The corresponding sequences of the second C_σ isotype are given below (vertical lines indicate identity). The second C_σ isotype is composed of two overlapping clones [i.e. clone σ12 (nt 9–650) and clone σ50 (nt 42–670)]. Gaps are introduced to maximize alignments. The stop codon TAA and the polyadenylation signal sequences (in italics) are underlined. Numbering of residues and nucleotides starts at the first amino acid and nucleotide of the constant region, respectively. The *RsaI* site (GTAC) common to $C_\sigma1$ and $C_\sigma2$ is double underlined.

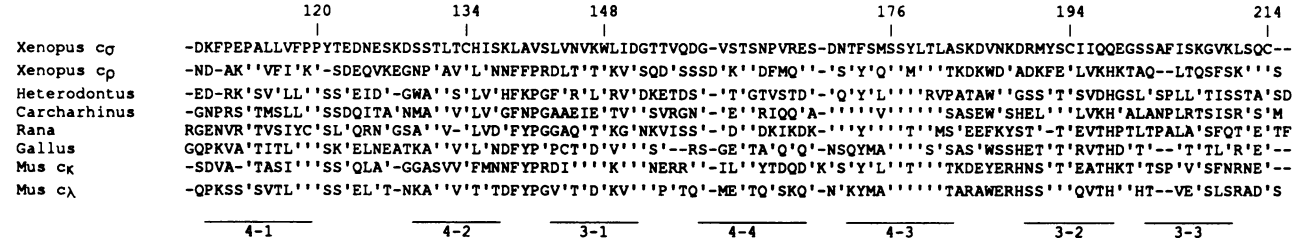

Fig. 3. Sequence alignment of the constant regions of IgL types from *Xenopus* and mammals. Protein sequence (in the one-letter code) from *Xenopus* $C_\sigma1$ (Figure 2), *Xenopus* C_ϱ (Zezza *et al.*, 1988), *Heterodontus francisci* IgL (Shamblott and Litman, 1989), *Carcharhinus plumbeus* IgL (Schluter *et al.*, 1989), *Rana catesbeiana* IgL (Mikoryak and Steiner, 1988), *Gallus* IgL (Reynaud *et al.*, 1985) and murine C_κ and C_λ (Kabat *et al.*, 1987) were aligned as described in Materials and methods. Areas of β-pleated sheet are indicated below the sequences and numbered according to Edmundson *et al.* (1975). Common residues are indicated by ditto, and some invariant residues numbered according to Kabat *et al.* (1987); gaps are introduced to maximize sequence similarity. C_σ has 29% identical residues with C_ϱ, 30% with *Heterodontus* C_L, 32% with *Carcharhinus* C_L, 28% with *Rana* C_L, 33% with chicken C_L, 31% with murine C_κ and 32% with murine C_λ.

leader peptides and in CDR1 and CDR2 (Figure 4D). The only differences are at the beginning and at the end of CDR3. Apparently only three $V_\sigma2$ elements which differ in residues 89 and 91 (clones σ10, σ29, σ117 and σ30; clones σ35 and σ120; clones σ110 and σ125) are expressed in these cDNA clones. Genomic Southern blots show that there are only a few $V_\sigma2$ elements (presumably <4; see below).

Junctional diversification and junctional imprecision in V_σ regions

In $V_\sigma2$ regions, residue 96 is the most variable in CDR3 and it is flanked by two invariant residues, Gly95 and Val97 (Figure 4D). Inspection of the nucleotide sequence provides evidence for junctional diversification during V_L-J_L rearrangements. All residues observed at position 96 in $V_\sigma2$ can be explained by the shift in the joining at the 3′ and 5′ ends of the two (hypothetical) V_L and J_L elements, respectively (Figure 6). In $V_\sigma2$ regions, the same $V_\sigma2$ elements are rearranged to $J_\sigma1$ or $J_\sigma2$ (e.g. clones σ29 and σ35). In $V_\sigma1$ elements, the CDR3 is more diverse. Some motifs at the end of CDR3 (e.g. A-W-V, A-Y-V) can also be explained by joining imprecision (Figure 6A).

In five of 12 $V_\sigma1^+$ cDNA clones, there is a frameshift at the end of the sequence encoding CDR3 (Figure 6). One (in clones σ102, σ103, σ111) or two (in clone σ21) nucleotides are lacking at the critical codons for CDR3 residues. These non-functional V_σ elements are rearranged to $J_\sigma1$ (clone σ103) or to $J_\sigma2$. The same V_σ element is abortively (clone σ21) or functionally (clone σ40) rearranged to the same J_σ element. In abortively rearranged V_σ regions, the sequences encoding residues 95–97 in functional V_σ regions (i.e. GCA TAT GTC or GCA TGG GTT) are retained in another reading frame (Figure 6).

Evidence for two Xenopus C_σ isotypes and multiple V_σ elements

We performed genomic Southern blot analyses with erythrocyte DNA from various individual *X.laevis* using probes specific for the C_σ region or for the V_σ/C_σ regions. Figure 7 shows that in all individuals tested, the C_σ probe [i.e. a 176 nt fragment of the constant region, which is virtually identical between $C_\sigma1$ and $C_\sigma2$ (see Figure 2)] hybridized to two to three restriction fragments. Using cDNA fragments derived from the 3′ untranslated region of $C_\sigma1$

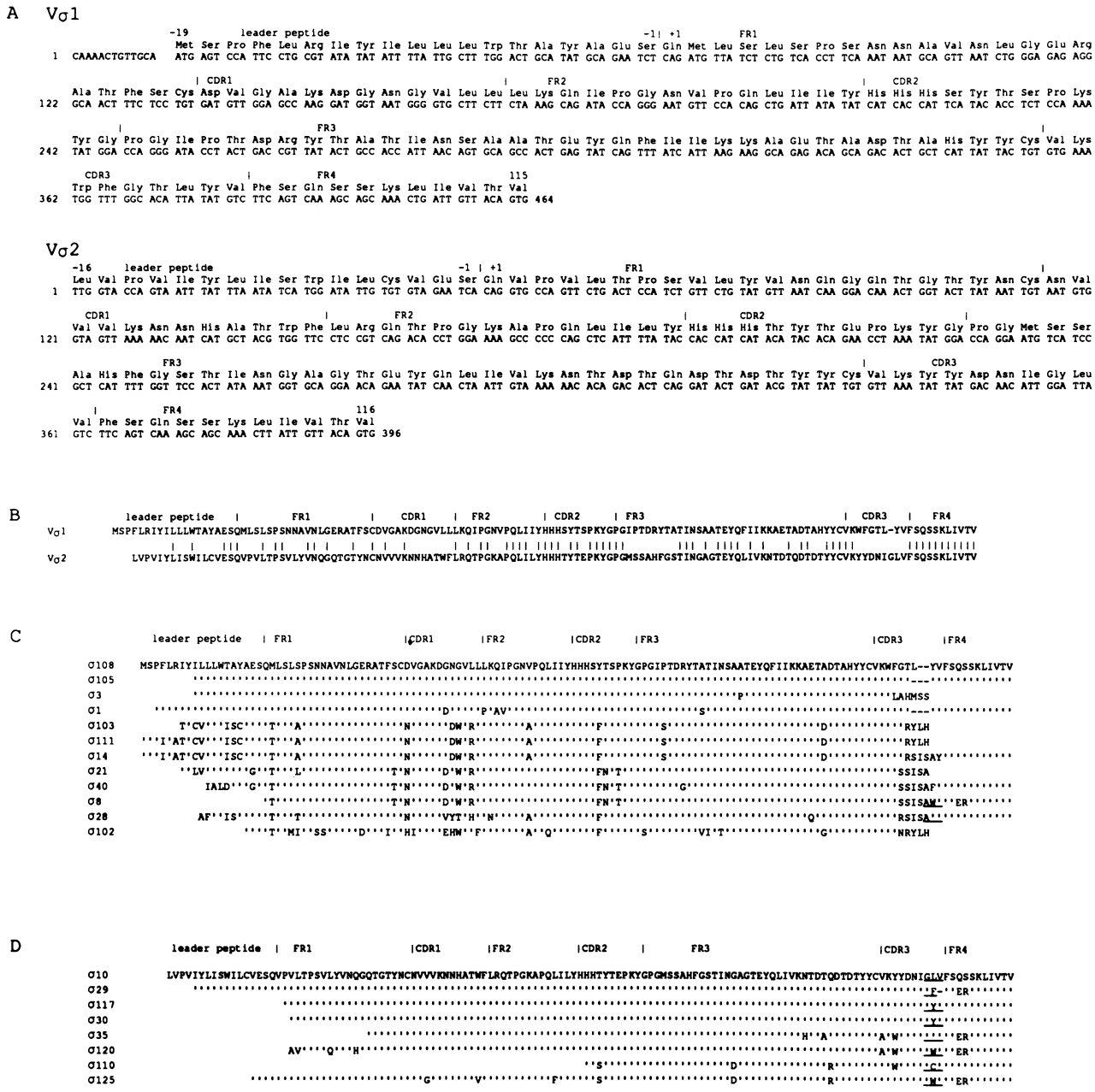

Fig. 4. Sequence of $V_\sigma1$ (A) and $V_\sigma2$ (B) regions. (**A**) Prototype $V_\sigma1$ sequence is from clone $\sigma108$ (see C) and from clone $\sigma10$ (see D). The beginning of the V_L region is positioned at Gln(+1) which implies a FR1 of 22 residues. Residues are numbered from Gln1 onward; nucleotides are numbered from the beginning of the respective cDNA clone. The boundaries between leader peptides, FRs and CDRs are according to Kabat *et al.* (1987). (**B**) Alignment of the $V_\sigma1$ and $V_\sigma2$ prototype protein sequences. Common residues are marked with dashes. (**C**) and (**D**) alignment of $V_\sigma1$ and $V_\sigma2$ regions. Only the protein sequence (in the one-letter code) is given. Boundaries of leader peptides, FRs and CDRs are as in (A). In the case of abortive rearrangements, the protein sequence is indicated until residue 95. Gaps are introduced to maximize alignments. Residues at the end of CDR3 which can be explained by junctional diversification (see Figure 6) are underlined.

and $C_\sigma2$ (i.e. an *RsaI*−*Eco*RI fragment of ~250 nt, see Figure 2), genomic restriction fragments of ~6 kb (or 5.5 kb) and ~3 kb, respectively, hybridized to the two probes. Therefore, the $C_\sigma1$ gene is located on a fragment of 6 kb or 5.5 kb depending on the individual animal, while the $C_\sigma2$ gene is located in the 3 kb fragment. All individuals tested possess both genes, which are therefore two C_σ isotypes. In contrast, multiple restriction fragments react with a probe containing $V_\sigma1-C_\sigma1$ and $V_\sigma2-C_\sigma1$ (Figure 7). $V_\sigma1$ elements are more abundant than $V_\sigma2$ elements (<15 DNA fragments versus <4 fragments in *PstI*−*Hin*dIII digested DNA from diploid animals). Similar patterns of restriction fragment length polymorphism (RFLP) are obtained in different outbred individuals (e.g. in animals 5, 6, 7, and animals 1 and 3, or 4 and 5, when probed with $V_\sigma1/C_\sigma1$ or $V_\sigma2/C_\sigma1$, respectively). Presumably, these unrelated individuals have common V_L pools.

Discussion

Both C_σ and V_σ reflect particular evolutionary pathways

The data presented in this paper illustrate two apparently conflicting findings. The sequence divergence between two *Xenopus* C_L regions, C_ϱ and C_σ is large, reflecting a rapid change or a long time since they separated; yet the expressed

Table I. Critical residues in *Xenopus* Vσ regions

V_L-V_H	residues	Vσ1	Vσ2
residues buried between the β-sheets			
4-4	*L M V*	L	P
6-6	*Q E L*	L	L
13-12	*G A V M I K*	L	V
19-18	*L V H I*	A	G
21-20	*L I M*	E	Y
23-22	*C*	C	C
25-24	*G A S T V*	V	V
33-34	*V I L M*	V	T
35-36	*W Y*	L	E
37-38	*Q R K H E*	K	R
47-48	*V L I W Y F E*	I	I
48-49	*I G A L M V*	I	L
64-69	*G A V I M F S L*	I	I
71-78	*A L F Y R V*	Y	Y
73-80	*L M N*	E	L
75-82	*L I M*	I	V
84-88	*G A S*	A	D
86-90	*Y*	Y	Y
88-92	*C*	C	C
99-104	*G*	S	S
101-106	*G*	S	S
102-107	*S T*	S	S
104-109	*V L*	L	L
106-111	*V L I*	V	V
residues in turns			
16-15	*G E*	G	G
41-42	*G R A*	G	G
61-66	*R K*	R	K
62-67	*V F I*	Y	Y
78-82c	*V L M I*	A	T
82-86	*D E*	D	D
residues between V_L and V_H			
36-37	*Y* *V F I*	L	L
38-39	*Q* *Q D E H*	Q	Q
44-45	*P* *L*	P	P
46-47	*L P R* *W*	L	L
87-91	*F Y* *F Y*	Y	Y
89-93	*A Q* *A L T S*	V	V
98-103	*F* *W*	F	F

Critical residues for mammalian V_L (number at left) and V_H regions (number at right) are from Chothia *et al.* (1987) and are completed with data from *Xenopus* V_H regions (Hsu *et al.*, 1989; Haire *et al.*, 1990; our unpublished results). Conserved residues in mammalian V_L and V_H regions are in bold letters, conserved residues in *Xenopus* V_H families are in italics. Conserved residues in FR1 and FR3 (residues 64–88) have been determined with respect to their distance from Cys23 and Cys88, respectively. Residues in $V_\sigma 1$ and $V_\sigma 2$ which deviate from those found in other V elements are underlined.

V_L elements are of a quite limited diversity. One way to resolve this apparent paradox is to assume that V_L and C_L genes have been subjected to very different evolutionary constraints.

C_σ, first described in the present paper, is as poorly related to C_ϱ as it is to the constant regions of sharks, *Rana catesbeiana*, chicken and mice (see Figure 3). This suggests that the $C_\varrho - C_\sigma$ divergence preceded speciation. Conserved residues are in the area of C_L that interacts with C_H, a result that underscores the constraints of contact residues during evolutionary processes in heterodimeric molecules. A detailed discussion of the various IgL types and their relation to other constant regions of members of the Ig gene superfamily will be presented elsewhere.

V_σ regions possess some residues within the β-pleated sheet of the Ig fold which deviate from those considered to be conserved, e.g. residues 35, 47, 99, 101 (Table I). Similarly, some *Xenopus* V_H families also have uncommon residues at the usually invariant Trp36 and Trp47 (Haire

```
          CDR3  |      FR4
             97 | 98
Jσ1  -CA TGG GTT TTC AGT GAA AGG AGC AAA CTG ATT GTT ACA GTG
Jσ2  -'' 'AT ''C ''' ''' C'' ''C ''' ''' ''' ''' ''' ''' '''

Jσ1       W  V  F  S  E  R  S  K  L  I  V  T  V
Jσ2       Y  '  '  '  Q  S  '  '  '  '  '  '  '

Jϱ1  --G TAC ACT TTT GGT AAA GGA ACC AGG GTA GAG ATA AAA CGA
Jϱ2  --T 'T' ''' ''' ''' ''' ''' ''' ''' ''' ''' C'T ''' '''

Jϱ1       Y  T  F  G  K  G  T  R  V  E  I  K  R
Jϱ2       L  T  '  '  '  '  '  '  '  '  L  '  '

Jκ1  --G TGG ACG TTC GGT GGA GGC ACC AAG CTG GAA ATC AAA CGT AAG
Jκ1       W  T  F  G  G  G  T  K  L  E  I  K  R  K

Jλ1  --C TGG GTG TTC GGT GGA GGA ACC AAA CTC ACT GTC CTA GGT
Jλ1       W  V  F  G  G  G  T  K  L  T  V  L  G
```

Fig. 5. Alignment of putative *Xenopus* J_L and murine J_L sequences. Codons of V_σ regions from residue 96 to the putative end of FR4 are aligned. $J_\sigma 1$ is from clone σ108, $J_\sigma 2$ from clone σ8. The operationally named $J_\varrho 1$ and $J_\varrho 2$ encode the ends of CDR3 and FR4 which is associated with *Xenopus* C_ϱ (Zezza *et al.*, submitted). The 5′ ends of J_σ and J_ϱ are not known from the cDNA clones. The sequences for murine $J_\kappa 1$ and $J_\lambda 1$ are from Kabat *et al.* (1987). Residues encoded by J_σ that are in common with $J_\kappa 1$ and with $J_\lambda 1$ are in italics and boxed, respectively. The nucleotide in codon Val97 which is diagnostic for $J_\sigma 1$ (i.e. T) or $J_\sigma 2$ (i.e. C) is underlined.

```
A
Vσ    GGA TTC G--        Vσ    GGA TGC G--
Jσ1   -CA TGG GTT        Jσ2   -CA TAT GTC

          A  W  V                  A  Y  V
σ8        GCA TGG GTT    σ28       GCA TAT GTC
          G  W  V                  G  Y  V
σ125      GGA TGG GTT    σ117      GGA TAT GTC
          G  L  V                  G  C  V
σ35       GGA TTG GTT    σ110      GGA TGC GTC
          G  F  -
σ29       GGA TTT ---

B
Vσ1
σ108  GTG AAA TGG TTT GGC ACA          TTA TAT GTC
σ105  --- --- --- --- --- ---             --- ---
σ3    --- --- --- --* --- ---             --- ---
σ1    --- --- --- --- --- ---             --- ---
σ111  --- --- --- --- --- -G- TAT CTG CAT ATG TC*
σ14   --- --- --- --- --- -GC ATA TCT GCA --- ---
σ21   --- --- --- --- A-- -GC ATA TCA GC- --- C**
σ40   --- --- --- --- A-- -GC ATA TCA GC- -TC ---
σ8    --- --- --- --- A-- -GC ATA TCA GC- -GG --T
σ28   --- --- --- --- A-G -GC ATA TCA GC- --- ---
σ102  --- --- --- --- AA- -G- TAT CTG CAT ATG TC*
σ103  --- --- --- --- --- -G- TAT CTG CAT GGG T-*

Vσ2
σ10   GTT AAA TAT TAT GAC AAC ATT GGA TTA GTC
σ29   --- --- --- --- --- --- --- --- --T
σ117  --- --- --- --- --- --- --- --- -AT ---
σ30   --- --- --- --- --- --- --- --- -AT ---
σ120  -C- --- -GG --- --- --- --- --- -GG --T
σ110  --- --- -GG --- --- --- --- --- -GC ---
σ125  --- --- --C --- --- --- --- --- -GG --T
```

Fig. 6. Diversification of CDR3 sequences. (**A**) Shown above are the putative V_L and the J_L elements (in italics), shown below are their respective contributions in the joining and the resulting residues, together with the cDNA clone that contains the sequence at the left. Note the marker nucleotide for $J_\sigma 1$ (i.e. Val97: GTT) and $J_\sigma 2$ (Val97: GTC). (**B**) Nucleotide sequences of CDR3 (i.e. between Cys88 and Phe98 (Kabat *et al.*, 1987) in functional and abortively rearranged $V_\sigma 1$ and $V_\sigma 2$ regions. The missing nucleotides that cause frameshift mutations are indicated by asterisks. A sequence which is common to functional and abortively rearranged V_σ regions but differs in the reading frame is underlined.

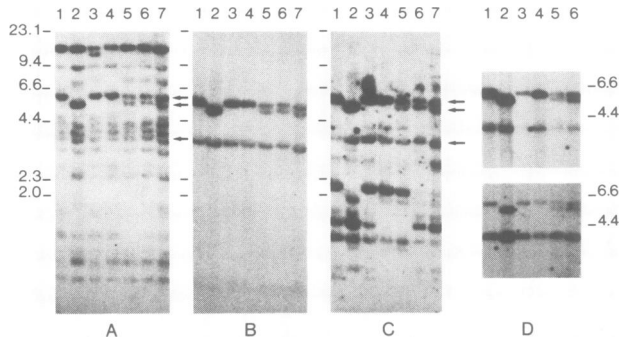

Fig. 7. Genomic Southern blot analysis of erythrocyte DNA from outbred animals. DNA was digested with *Hin*dIII−*Pst*I. The blots were hybridized with (**A**) $V_\sigma1-C_\sigma1$ probe (insert of clone σ1); (**B**) a C_σ probe (176 nt of the coding region); (**C**) $V_\sigma2-C_\sigma1$ probe (insert of clone σ10) or (**D**) with a probe specific for the untranslated 3′ end of $C_\sigma1$ (upper panel) and of $C_\sigma2$ (lower panel). Markers (in kb) are indicated at the margins. The positions of fragments of the C_σ genes in panels A and C are indicated by arrows.

et al., 1990 and our unpublished results). Most of the critical residues of V_L regions that interact in the β-pleated sheet of the Ig fold or with V_H are conserved (Table I). The implication of the serine or arginine residues at position 99 and 101 on the Ig fold remains to be elucidated. When $V_\sigma1$ sequences were compared with a database of amino acid sequences (see Materials and methods) the best scores were obtained with $V_\varkappa$ regions; V_λ regions and V regions of other members of the Ig gene superfamily were less concordant with V_σ regions. Thus, $V_\sigma1$ regions are related to V_L elements but they differ from them in length, presumably due either to an unusually long loop in the CDR2 or to the length of FR3. In contrast, a similar search with $V_\sigma2$ regions showed that the carboxy-terminal half of $V_\sigma2$ regions is more similar to V_H regions than it is to V_L regions. The evolutionary significance of these findings remains to be determined. The V_σ subgroups described here have FR1 sequences that differ from the IgL chain N-terminal sequences obtained in non-immune serum and anti-DNP specific antibodies (Wang *et al.*, 1978; Brandt *et al.*, 1980). V_σ regions have an N-terminal glutamine which presumably cyclizes to pyroglutamic acid, rendering V_σ proteins inaccessible to direct N-terminal sequence determination; therefore V_σ regions belong to the ~95% *Xenopus* IgL chains that have blocked N-terminal residues (Wang *et al.*, 1978). V_L families different from V_σ might constitute a part of the IgL chain repertoire.

IgL loci are similar in Xenopus and in mammals

Xenopus IgL loci appear to be organized like mammalian IgL loci. There are multiple $V_\sigma1$, a few $V_\sigma2$ elements and two C_σ regions (Figure 7), $C_\sigma1$ and $C_\sigma2$. In cDNA libraries from outbred animals, among >20 different cDNA clones, only two are derived from the $C_\sigma2$ gene. Presumably, it forms a transcription unit with $C_\sigma1$, but it is rarely used in putative alternative splicing events where V_LJ_L proximal elements are favoured (i.e. $C_\sigma1$). Our data suggest that the *Xenopus* IgL loci differ from those of sharks (Schluter *et al.*, 1989; Shamblott and Litman, 1989) and chickens (Reynaud *et al.*, 1985) which express only one light chain type. IgL and IgH loci (Schwager *et al.*, 1988b) have a similar organization in *Xenopus* and thus have co-evolved as they did in other species. The V_σ regions described here are only associated with C_σ, and never with C_ϱ which has its own set of V_L elements (our unpublished observations). This is reminiscent of the expression of λ and ϰ light chains in mammals. Indeed, C_ϱ and C_σ are located on different linkage groups (our unpublished results).

Are $V_\sigma1$ and $V_\sigma2$ sequentially expressed during B cell development?

Diversity of CDR3 increases by rearranging V_L to J_σ elements. Yet $V_\sigma1$ elements are frequently abortively rearranged to J_L elements. A given V_L element might be functionally expressed as a result of another rearrangement event (see clones σ21 and σ40 or clones σ103 and σ111). In contrast, the few $V_\sigma2$ elements (see Figure 7) rearrange successfully, but they display a lower CDR3 diversity. We hypothesize that these differences in expression of $V_\sigma1$ and $V_\sigma2$ reflect a specific developmental programme in B cell development, in which $V_\sigma1$ elements are rearranged first on one and, in the case of abortive rearrangements, also on the second allele. Potentially, this process creates high diversity. Subsequently, in the absence of functional $V_\sigma1$ regions, $V_\sigma2$ rearrangements occur in a precise way but at the expense of increased diversification. Alternatively, $V_\sigma2$ elements are rearranged before $V_\sigma1$ elements and may constitute a primary B cell pool with a restricted heterogeneity; subsequent rearrangements of $V_\sigma1$ elements diversify V_L regions later in B cell development. This latter interpretation is consistent with findings in the larval and adult B cell compartments (our unpublished results): heterogeneity of antibody populations is acquired step-wise during larval and adult life (Hsu and Du Pasquier, 1984); presumably, it reflects successive waves of B cell development (Hadji-Azimi *et al.*, 1982, 1990).

Diversity of IgL populations and association with heavy chains

Mitogen-stimulated *Xenopus* spleen cells produce IgL chains which appear to be as heterogeneous as their mammalian counterparts (Schwager and Hadji-Azimi, 1984, 1985). Although $V_\sigma1$ and $V_\sigma2$ regions differ by many charged residues (Figure 4C and D), the data obtained from the cDNA clones may not fully account for the relatively heterogeneous IgL population. The data are unlikely to be biased by the method of identifying cDNA clones: clones for different IgL types and for all three IgH classes (IgM, IgX and IgY) were obtained from the same library and were associated with at least five different V_H families. Clones with the same features as the C_σ^+ cDNA clones described here have been isolated from other cDNA libraries (unpublished observation).

One IgL population (i.e. the ~27 kd protein) is preferentially associated with υ chains, while μ chains form IgM molecules with the 25 kd IgL chain population (Hadji-Azimi, 1975; Hsu *et al.*, 1985). Consequently, based on the results of the immunoscreening obtained in the present study, σ and ϱ chains prefer to form Ig molecules with μ and υ chains, respectively. Northern blot analyses, however, reveal that, even in the absence of significant amounts of υ mRNA, both ϱ and σ are transcribed in adult splenic B cells (Figure 1). It seems therefore that IgH gene activity does not regulate the expression of IgL types.

The IgL loci may limit antibody diversity in Xenopus

The analysis of antibody diversity by isoelectric focusing (Du Pasquier and Wabl, 1978) and of the inheritance of idiotypes

(Brandt *et al.*, 1980) revealed a relatively limited antibody repertoire in *Xenopus* as compared with that in mammals. Due to a large number of V_H elements, putative D_H elements and J_H elements, the three largest V_H families have a high potential to create diversity (Schwager *et al.*, 1989). Sharing of CDRs within V_H families, distinct promoter elements which may influence V_H usage and frequent pseudogenes within one large family only partially explained the restricted antibody diversity. The data presented here suggest that the expressed V_σ pool restricts antibody diversity. If the complete V_L pool were similarly limited, IgL chains would be the limiting factor for antibody diversity in *Xenopus*. In functional Ig molecules, the IgL chain appears to be less crucial in the specific antigen–antibody interaction (Ward *et al.*, 1989) and may therefore be relatively invariant. Consequently, in *Xenopus* a few expressed V_L elements might limit the structural, but not necessarily the functional, diversity of antibody molecules. Further analysis of the genomic V_L elements are necessary to clarify whether limited V_L usage and complexity, or distinct regulatory elements account for this restriction.

Materials and methods

Construction of cDNA libraries and isolation of cDNA clones

cDNA libraries were constructed from poly(A)$^+$ RNA which was isolated from mitogen-stimulated spleen cells from 10 outbred adult *Xenopus* (see also Schwager *et al.*, 1988a), or from total RNA (40 μg) isolated from spleens of outbred *Xenopus*; cDNA was synthesized using a commercial cDNA synthesis kit (Pharmacia) and cloned into the lambda ZapII vector system (Stratagene). Immunoscreening with rabbit anti-*Xenopus* IgM, plaque hybridization and purification were done as described previously (Schwager *et al.*, 1988a,b). Anti-IgM antisera were obtained by injecting rabbits twice with 50 μg of periodate-treated *Xenopus* IgM (Hadji-Azimi and Schwager,1980). The IgG fraction obtained from hyperimmunized rabbits was prepared and used at 500 μg/ml in the immunoscreening. This antiserum reacts with *Xenopus* μ and IgL chains as determined by Western blotting. Recombinant cDNA was *in vivo* excised with the M13 VCS helper phage following the protocol suggested by the manufacturer (Stratagene).

Analysis of the cDNA clones

Double-stranded recombinant plasmid DNA was sequenced on both strands as described previously (Schwager *et al.*, 1989) using universal forward and reverse primers and synthetic oligonucleotides (15–24mer, kindly provided by Dr H.R.Kiefer, Basel Institute for Immunology), for J_L elements or C_L sequences.

RNA and DNA blot analysis

Adult spleen cells were stimulated with PWM (Gibco, 1/20 final concentration) or with LPS (Sigma, 150 μg/ml) in Click's medium for 6–8 days as described (Schwager and Hadji-Azimi, 1984). RNA was extracted from different tissues or from cultured cells, denatured, electrophoresed and blotted to Duralon nylon membranes (Stratagene) (Schwager *et al.*, 1988a). Genomic Southern blot analysis and conditions for hybridization to random-primed ^{32}P-labelled DNA probes were similar to those described (Schwager *et al.*, 1988b), except that the hybridization was performed in 6 × SSC, 1 × Denhardt's solution, 0.1% SDS, 10 μg/ml *Escherichia coli* DNA, 8 μg/ml poly(A) and poly(C) at 65°C for 16 h. A probe common to $C_\sigma1$ and $C_\sigma2$ extends from nt 73 to nt 247 in clone σ6. Probes specific for $C_\sigma1$ and $C_\sigma2$ were obtained from *Rsa*I–*Eco*RI fragments of clone $C_\sigma6$ and $C_\sigma12$, respectively (see Figure 2). In DNA blots, filters were washed with 0.1 × SSC, 0.1% SDS at 68°C for 2 × 20 min and exposed for 1–3 days.

Computer analysis of nucleotide sequences

DNA sequences were analysed with a computer program written by Charley Steinberg (Basel Institute for Immunology), which is based on the algorithm Needleman–Wuntsch. DNA sequences were compared with published sequences (GenBank release 63); their deduced protein sequences were derived by translating all coding sequences contained in the GenBank (release 63). (This translation program was also written by Charley Steinberg.) Proteins were aligned using the Dayhoff weight table (Dayhoff *et al.*, 1983). Gap penalty was 10 and 6 for nucleotide and amino acid sequence alignments, respectively.

Acknowledgements

We thank Dr H.R.Kiefer's laboratory for synthesizing oligonucleotides, Anne Marcuz for help in Southern blot analysis and Charley Steinberg for invaluable assistance in using his computer programs. Drs Antonio de la Hera, Louis Du Pasquier, Jim Kaufmann and Charley Steinberg are acknowledged for critical reading of the manuscript and useful comments. We thank Hans Spalinger and Bea Pfeifer for photographic work.

The Basel Institute for Immunology was founded and is supported by F.Hoffmann-La Roche Ltd, Basel, Switzerland.

References

Alt,F.W., Blackwell,T.K. and Yancopoulos,G.D. (1987) *Science*, **238**, 1079–1087.

Brandt,D.C., Griessen,M., Du Pasquier,L. and Jaton,J.-C. (1980) *Eur. J. Immunol.*, **10**, 731–736.

Chothia,C., Boswell,D.R. and Lesk,A.M. (1988) *EMBO J.*, **7**, 3745–3755.

Dayhoff,M.O., Barker,W. and Hunt,L.T. (1983) *Methods Enzymol.*, **91**, 524–545.

Du Pasquier,L. (1982) *Nature*, **296**, 311–313.

Du Pasquier,L. (1989) In Paul,W.E. (ed.), *Fundamental Immunology*. Raven Press, New York, pp. 139–165.

Du Pasquier,L. and Wabl,M. (1978) *Eur. J. Immunol.*, **8**, 428–435.

Du Pasquier,L., Schwager,J. and Flajnik,M.F. (1989) *Annu. Rev. Immunol.*, **7**, 251–275.

Edmundson,A.B., Ely,K.R., Abola,E.E., Schiffer,M. and Panagiotopoulos, N. (1975) *Biochemistry*, **14**, 3953–3961.

Hadji-Azimi,I. (1975) *Immunology*, **28**, 419–429.

Hadji-Azimi,I. and Schwager,J. (1980) *Cell Immunol.*, **53**, 389–394.

Hadji-Azimi,I., Schwager,J. and Thiébaud,C.H. (1982) *Dev. Biol.*, **90**, 253–258.

Hadji-Azimi,I., Coosemans,V. and Canicatti,C. (1990) *Dev. Comp. Immunol.*, **14**, 69–84.

Haire,R.N., Amemiya,C.T., Suzuki,D. and Litman,G.W. (1990) *J. Exp. Med.*, **171**, 1721–1737.

Hasemann,C.A. and Capra,J.D. (1989) In Paul,W.E. (ed.), *Fundamental Immunology*. Raven Press, New York, pp. 209–233.

Hinds,K.R. and Litman,G.W. (1986) *Nature*, **320**, 546–549.

Hsu,E. and Du Pasquier,L. (1984) *Mol. Immunol.*, **21**, 257–269.

Hsu,E. and Du Pasquier,L. (1984) *Differentiation*, **28**, 116–122.

Hsu,E., Flajnik,M.F. and Du Pasquier,L. (1985) *J. Immunol.*, **135**, 1998–2204.

Hsu,E., Schwager,J. and Alt,F.W. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 8010–8014.

Kabat,E.A., Wu,T.T., Reid-Miller,M., Perry,H.M. and Gottesman,K.S. (1987) *Sequences and Proteins of Immunological Interest*. 4th edn. HHS, National Institutes of Health, Bethesda, MD.

Mikoryak,C.A. and Steiner,L.A. (1988) *Mol. Immunol.*, **25**, 695–703.

Reynaud,C.A., Anquez,V., Dahan,A. and Weill,J.-C. (1985) *Cell*, **40**, 283–291.

Reynaud,C.-A., Dahan,A., Anquez,V. and Weill,J.-C. (1989) *Cell*, **59**, 171–183.

Schluter,S.F., Hohman,V.S., Edmundson,A.B. and Marchalonis,J.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9961–9965.

Schwager,J. (1983) PhD thesis work no. 2105, University of Geneva.

Schwager,J. and Hadji-Azimi,I. (1984) *Differentiation*, **27**, 182–188.

Schwager,J. and Hadji-Azimi,I. (1985) *Differentiation*, **30**, 29–34.

Schwager,J., Mikoryak,C.A. and Steiner,L.A. (1988a) *Proc. Natl. Acad. Sci. USA*, **85**, 2245–2249.

Schwager,J., Grossberger,D. and Du Pasquier,L. (1988b) *EMBO J.*, **7**, 2409–2415.

Schwager,J., Bürckert,N., Courtet,M. and Du Pasquier,L. (1989) *EMBO J.*, **8**, 2289–3001.

Shamblott,M.J. and Litman,G.W. (1989) *EMBO J.*, **8**, 3733–3739.

Tonegawa,S. (1983) *Nature*, **302**, 575–581.

Wang,A.C., Tung,E., Fudenberg.H.H. and Hadji-Azimi,I. (1978) *J. Immunogenet.*, **5**, 355–364.

Ward,E.S., Güssow,D., Griffiths,A.D., Jones,P.T. and Winter,G. (1989) *Nature*, **341**, 544–546.

Zezza,D.J., Mikoryak,C.A., Schwager,J. and Steiner,L.A. (1988) *Fed. Proc., (abstract)*.

Received on November 26, 1990