

Lars Hellman

Department of Immunology,
University of Uppsala, Biomedical
Center, Uppsala

Characterization of four novel ϵ chain mRNA and a comparative analysis of genes for immunoglobulin E in rodents and man*

Questa fotocopia viene fornita
per motivi strettamente personali
di studio e/o ricerca

The nucleotide sequence of the 3' region of the ϵ chain gene for human IgE is presented. A comparison of the entire region from 5' of exon C1 to the M2 exon of the mouse, rat and human ϵ chain genes shows that the overall structure of the ϵ chain gene have changed only minimally during the 60–70 million years of evolutionary separation between rodents and man.

We have previously shown that a number of rearrangements larger than 10 bp have relatively recently occurred in the C4/M1 intron of the rat or the mouse ϵ chain genes. A majority of these rearrangements were found within or in close proximity to repetitive sequences of Z-DNA-forming potential (CA dinucleotide repeats). The C4/M1 intron has evolved very rapidly, to such an extent that no apparent homology can be detected between rodents and man. Only remnants of the repetitive sequences are present in man, supporting the theory that repetitive sequences having Z-DNA-forming properties may play a role in the evolution of the eucaryote genome by promoting recombinations, leading to a rapid evolutionary drift of sequences in close proximity to these repeats.

We report here the characterization of the membrane domains of human IgE and four novel mRNA transcribed from the human ϵ chain locus. The primary structures have been determined by polymerase chain reaction cloning and nucleotide sequence analysis. All five mRNA contain the C3 domain and the membrane exon 2 (M2). Due to frame shifts caused by novel splice sites or novel splice-site combinations, the proteins encoded by three out of these four novel mRNA differ in their carboxy-terminal end from the classical secreted or membrane-bound immunoglobulins. Northern blot analysis shows significant levels of at least three out of these four novel mRNA in an IgE-producing human cell line.

One of the mRNA encodes a transmembrane-like structure which has characters in common with the transmembrane region of the CD3 components of the T cell receptor complex (CD3 γ , δ and ϵ). This indicates that IgE-producing B cells possibly have two separate signal-transducing systems.

A comparison of the classical membrane anchoring domain of the human ϵ chain with a panel of immunoglobulin membrane domains from fish to higher mammals is presented. A tyrosine and a glutamine residue is found to be conserved between all cytoplasmic domains of all post-switch immunoglobulin classes indicating a functional conservation of these amino acid residues. Suggested functions for these residues are their participation in signal transduction or in linkage to the cytoskeleton for internalization of antibody-antigen complexes.

1 Introduction

IgE has attracted a major interest since its discovery in 1967 by Johansson and Bennich [1] and Ishizaka and Ishizaka [2], due to the central role it plays in atopic allergies. Although the normal function of IgE is not known, it has been suggested that IgE could be of importance for the

immune defense against parasitic infections [3]. A comparison of the corresponding peptide regions of human and rodent Ig shows that the ϵ chain has the lowest degree of conservation of all five Ig classes. Since the high-affinity interaction between the IgE antibody and its high-affinity receptor on mast cells and basophilic leukocytes is maintained in both rodents and man, this is strongly indicative of a co-evolution of the antibody and its receptor and an evolutionary advantage of maintaining the IgE mast cell system.

The nucleotide sequence has previously been determined for the region containing the four constant region exons (C1 to C4) of the ϵ chain gene of IgE and the region containing the tentative position of the two exons (M1 and M2) encoding the membrane anchoring tail of the membrane-bound form of IgE in mouse, rat and man [4–8]. By a comparative analysis of the complete region from 200 bp upstream of exon C1 to a few hundred bp downstream of the coding region of the M2 exon of the ϵ chain genes in

[I 10391]

* This investigation was supported by grants from the Swedish Natural Sciences Research Council (B-BU 9400-301) and the Knut and Alice Wallenberg Foundation.

Correspondence: Lars Hellman, Department of Immunology, University of Uppsala, The Biomedical Center, Box 582 S-751 23 Uppsala, Sweden

Abbreviations: aa: Amino acids nc: Noncoding region

Key words: IgE / Transcription / Rodent / Human

mouse and rat, we have demonstrated that the positions of all larger insertions and deletions in this chromosomal region are, to a large extent, localized to repetitive sequences containing repeated CA dinucleotides [9]. Such repetitive sequences have been found, *in vitro* and possibly *in vivo*, to adopt Z-DNA conformation. The localization of the rearrangements to the repetitive sequences were striking, implicating a role of these sequences in gene rearrangements and the evolutionary drift of the eucaryote genome. This prompted an analysis of the corresponding region in man. Primates and rodents separated during evolution approximately 60–70 million years ago whereas the evolutionary separation of mouse and rat occurred approximately 10–20 million years ago.

Except for the sterile transcripts preceding the isotype switch [10–12], the mRNA encoding the classical secreted and membrane-bound antibodies have long been considered the only mRNA transcribed from an Ig locus. A few examples have been described where two or sometimes three separate transcripts are present for the membrane Ig [13]. These have been shown to differ only in the position of their polyadenylation sites [13]. Here we report the characterization of the classical membrane IgE and four novel mRNA originating from the human ϵ chain locus. Three of these encode proteins which clearly differ from the classical secreted and membrane-bound IgE. One of the novel proteins contains a potential transmembrane region which shows similarities to the CD3 components of the T cell receptor. This indicates that IgE-producing B cells potentially have a second signal-transducing system separate from the classical membrane-bound Ig. Recently a study was published describing the structure of the classical membrane form of human IgE and two additional transcripts encoded from the ϵ chain locus [14]. One of these mRNA is identical to one of the transcripts we describe in this communication and the second encodes a novel variant of the classical membrane anchoring IgE with an additional 51 amino acids (aa) positioned between the C4 domain and the transmembrane region. The structure of the present seven different mRNA transcribed from the ϵ chain locus is compared and their possible biological function and significance will be discussed.

The transmembrane Ig is, together with the accessory signal-transducing molecules encoded by the genes mb-1 and B-29, essential for antigen-dependent signal transduction and antigen presentation by the B cell. A comparative analysis of membrane-spanning and cytoplasmic domains of a large panel of Ig genes from fish to man is presented in an attempt to define aa residues of importance for these two important events in B cell development.

2 Materials and methods

2.1 Subcloning and nucleotide sequence analysis

Following a detailed restriction enzyme mapping of the cosmid clone CosIg10 [15], a Hind III/Xba I fragment of approximately 4.8 kb in size containing the four constant region exons and 1.5 kb of 3' sequences of the ϵ chain gene was excised from the cosmid by restriction enzyme cleavage and purified on a low-melting-point agarose gel. The

fragment was ligated into pGEM2. A 976-bp Sac I fragment and a 510-bp Nhe I/Pst I fragment from the 3' end of the ϵ chain gene was excised from the cosmid clone and subcloned into pGEM2. Detailed restriction maps were generated for the subclones and fragments of 200–400 bp in size were subcloned for nucleotide sequence analysis by the method of Sanger et al. [16]. All sequences depicted in Fig. 1 were determined for both directions of the DNA molecule. The cosmid clone was a kind gift from Dr. T. Rabbitts.

2.2 Northern blot analysis

Total cellular RNA was purified from the human IgE-producing B cell line U-266 as previously described [17]. The RNA was further purified by one cycle of oligo (dT) selection prior to separation on formaldehyde-containing 1.0% agarose gel [18]. The gels were blotted on Hybond C extra membranes (Amersham Int., Amersham, GB) [18]. The filters were hybridized with seven different probes originating from various regions of the human ϵ chain gene (see Fig. 6 for details). Probes were labeled by random priming (Amersham Multiprime). The filters were washed at medium stringency $2 \times$ SSC, 0.5% SDS at 65 °C.

2.3 Polymerase chain reaction (PCR) amplification

Single-stranded cDNA was prepared as follows: 5 μ g of poly A⁺ RNA was added to a solution containing 50 mM Tris-HCl (pH 7.6), 60 mM KCl, 10 mM MgCl₂, 1 mM DTT, 1 U/ μ l of RNasin, 50 μ g/ml of actinomycin D and 1 U/ml AMV reverse transcriptase, to a final volume of 40 μ l. After 2 h of incubation at 42 °C the RNA template was removed by addition of NaOH to a final concentration of 0.25 M and the reaction was continued at 65 °C for an additional 1 h. The sample was neutralized by the addition of 1/10 volume of 1 M Tris-HCl pH 7.5 and 5 M HCl in an amount equivalent to the NaOH added to the reaction tube. Two hundred and fifty microliters of a solution containing 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 0.1 M NaCl, and 0.5% SDS was added and the sample was extracted with phenol:chloroform (50:50). This was followed by chloroform extraction. The cDNA was ethanol precipitated, washed in 70% ethanol and dissolved in 20 μ l double-distilled water. Five microliters of this cDNA was used as template in a PCR amplification performed according to standard procedures [19].

3 Results

3.1 Nucleotide sequence analysis of the 3' region of the human ϵ chain gene

Three fragments originating from the ϵ chain gene of the cosmid clone Cos Ig10 [15] were subcloned into pGEM 2 for further analysis. The nucleotide sequence was determined for a region from the fourth constant region to approximately 800 bp downstream of the coding region of exon M2 (Fig. 1). The tentative positions of exons M1 and M2 were determined on the basis of homology to the rodent counterparts [4, 5].

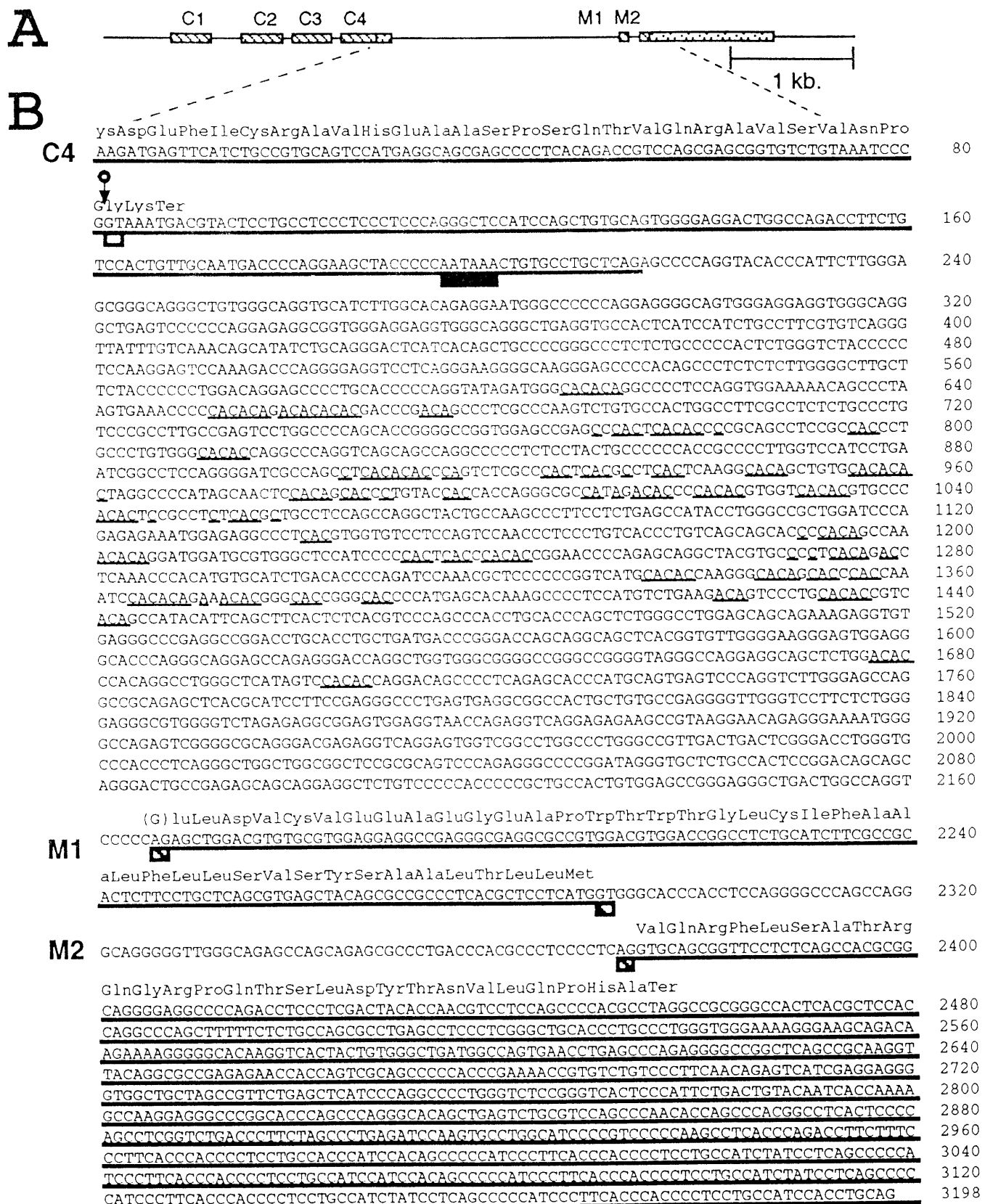


Figure 1. (A) shows a schematic map (to scale) of the human ϵ chain locus. The hatched boxes indicate the position of the coding regions and the dotted boxes the nc regions of all exons (C1-C4 = constant region exons 1 to 4; M1 and M2 = membrane domain exons 1 and 2). (B) shows the complete sequence of the region from C-terminal end of exon C4 to the region of exon M2. Exon sequences are marked by a thick black line below the corresponding sequence. Splice sites for the M1 and M2 exons are shown by hatched boxes below the corresponding sequence. The splice site in the C4 exon is marked by an empty box below, and an arrow above the actual site. The coding regions are indicated by the aa sequence above the corresponding nucleotide sequence. The polyadenylation signal is indicated by a black box below the actual sequence in the 3' end of exon C4. Thin lines below the sequence in the C4/M1 intron show the remnants of the CA repeats described in the Sect. 3.2.

3.2 Nucleotide sequence comparison of genes for the ϵ chain of IgE in rodents and man

The nucleotide sequence of the human ϵ chain gene was compared to the mouse and rat counterparts using a dot matrix computer program (Mac Molly-Compare, windows 25 nt. allowing 9 nonmatching nt; Fig. 2). Significant homologies were found only in and around all exons. The larger introns (C1/C2 and C4/M1) have diverged to such an extent that no apparent homologies can be found (Fig. 2).

Four large clusters of repeated CA dinucleotides are present in the C4/M1 intron of the rat ϵ chain gene (Fig. 3 A). No longer repeats, but multiple short stretches of CA-containing sequences, were found in the C4/M1 intron in man indicating that the CA repeats were present in the intron prior to the evolutionary separation of rodents and man (Fig. 1).

The size of the C4/M1 intron (from the poly A site) is 1762 bp in mouse and 1954 bp in man [9], while rat differs by approximately 475 bp compared to mouse (an intron size of 2237 bp) [9]. When these extra sequences were analyzed four insertions or duplications, larger than 10 bp, were found [9]. Two of these (22 bp and 90 bp) were found to be direct duplications or deletions inside a block of repetitive sequences (Fig. 3 A). One small duplication of 22 bp was found outside of the repetitive sequences (Fig. 3 A) [9]. One large insertion of 349 bp was found adjacent to two blocks of repetitive sequences (Fig. 3 A). No apparent homology to this 349-bp sequence was found to any other region of the rat ϵ chain gene, showing that this extra sequence did not originate from a duplication and thereby must be considered as a direct insertion [9] (Fig. 3 A).

An insertion of a highly AT-rich sequence has been found in the C1/C2 intron of the mouse [4, 5] (Fig. 3 A). This insertion of 366 bp is flanked by a 22-bp direct repeat resembling the situation found for many transposable elements [5]. No larger open reading frame is found in this

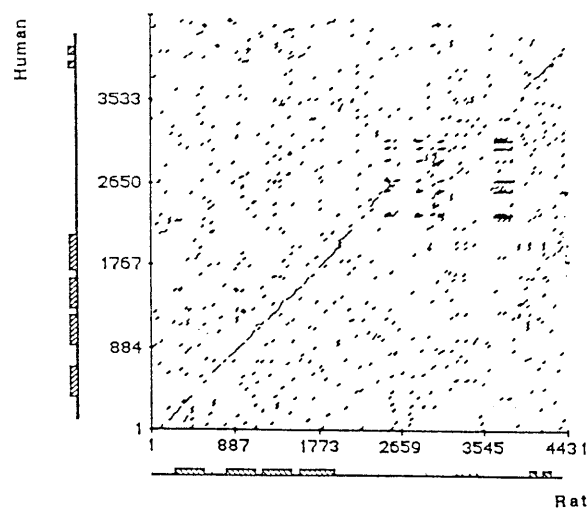


Figure 2. A dot matrix comparison of the human and rat ϵ chain genes. The exon-intron structure of the human and rat ϵ chain genes is shown on the Y and X axes of the figure, respectively. The test window in the dot matrix program Mac Molly Compare was 25 nucleotides long with a maximum of 9 mismatches allowed in order to give a line of homology.

insertion, indicating that this DNA element does not code for a protein. The origin and possible function of this sequence is not known. We have shown that this sequence is present in a few hundred copies in the mouse genome (Steen et al., unpublished results) and is transcribed at a low level as an approximately 400-nt transcript (unpublished results Hellman et al.).

3.3 PCR amplification to determine the structure of the 3' region of the membrane form of the human ϵ chain mRNA

To confirm the predicted structure of the membrane exons and to determine the exact splice sites for the membrane forms, two oligonucleotide primers were synthesized for PCR amplification of the 3' end of these mRNA. These oligonucleotides were complementary to sequences from the 3' region of exon C3 and a sequence starting 35 bp downstream of the predicted stop codon in exon M2. Single-stranded cDNA, prepared from mRNA of the human IgE-producing myeloma U-266 [20, 21], was used as starting material in a PCR amplification. Two main bands of approximately 500 bp and 750 bp were detected. DNA from both of these bands was purified and subcloned into pGEM2 for further analysis. Ten clones originating from each of the bands were characterized by nucleotide sequence analysis. Among these clones the mRNA encoding the classical membrane-bound Ig and four novel mRNA were found. Based on this sequence information we were

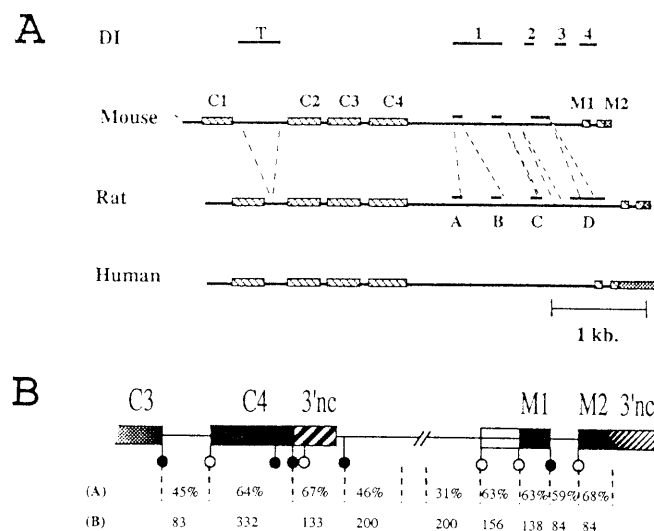


Figure 3. (A) shows a map of the regions analyzed by detailed nucleotide sequence determination for the ϵ chain genes of IgE in mouse, rat and man. The position of the transposon-like insertion in the mouse ϵ chain gene is indicated by a large T above the corresponding region in the line marked DI (deletions and insertions). The positions of the four larger insertions and deletions found in the C4-M1 intron described in [9] is shown by the numbers 1 to 4. The four large blocks of CA repeats in the rat C4/M1 intron are indicated by the letters A to D below the corresponding location in the figure of the rat ϵ chain locus. (B) The exons and the splice sites described in Fig. 4 are shown in the upper part of the figure for an easy identification of those regions the sequence conservation refers to. In row (A) the relative level of sequence conservation is shown in % below the corresponding region. In row (B) the total size of the regions used for the comparison is indicated by the number of base pairs.

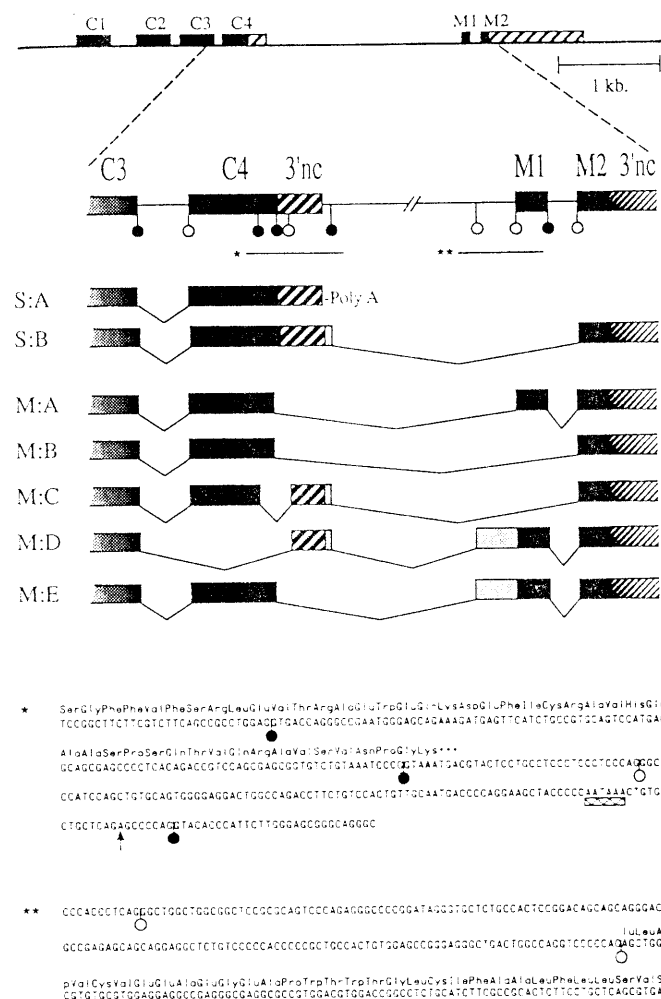


Figure 4. At the top of the figure a schematic map (to scale) of the human ϵ chain locus is presented. The black boxes indicate the position of the coding regions of exons (C1-C4 = constant region exons 1 to 4; M1 and M2 = membrane domain exons 1 and 2). The hatched boxes indicate the nc regions of the mRNA for the classical membrane and secreted IgE molecules. Below the schematic map we find an extended map of the exon regions where the splice sites are indicated by filled or open circles. Donor sites are indicated by filled circles and acceptor sites by open circles. The regions whose full sequence is shown at the bottom of the figure are indicated by a thin line below the detailed map of the locus, with one or two asterisks referring to the corresponding sequence in the lower part of the figure. In the central region of the figure we show the splicing pattern of the present seven different mRNA transcribed from the human ϵ chain gene. The boxes indicate the exons, their positions and their approximate sizes. The boxes in grey are the regions of the novel exons which are not present in the exons of the mRNA for the classical membrane or secreted IgE molecules. In the lower part of the figure the nucleotide sequence of the regions where we find the four novel splice sites described in the text are depicted. The sequence indicated by one asterisk is the 3' end of exon C4 where the carboxy-terminal end of the reading frame for the secreted IgE molecule is shown above the corresponding nt sequence. The sequence indicated by two asterisks is the 5' region of exon M1 with the novel 156-bp exon region present in the mRNA M:D and M:E. The aa sequence of the M1 exon is depicted above the corresponding nt sequence. This reading frame is utilized in the mRNA M:A. Donor splice sites are indicated by filled circles and acceptor sites by open circles. The polyadenylation signal is indicated by a box and the polyadenylation site by a black arrow below the corresponding sequence.

able to determine the exact positions of the exons M1 and M2 and the exon boundaries for four novel ϵ chain mRNA (Fig. 4).

3.4 Characterization of the classical membrane IgE and four novel mRNA transcribed from the ϵ chain locus

One of the novel mRNA, the S:B, encodes a secreted IgE molecule identical to the classical secreted IgE (Fig. 4). The difference between this mRNA and the earlier described mRNA for the secreted IgE, the S:A is located in the 3' non-coding (nc) region (Fig. 4). S:B has here an extra extension, originating from a novel donor site located 8 nt downstream of the poly A site of S:A. This donor site is ligated directly to the acceptor site of exon M2 resulting in an mRNA which is probably polyadenylated at a poly A site common to all M2-containing mRNA. Based on Northern blot analysis this poly A site is estimated to be located 1 kb downstream of the acceptor site for exon M2 (Fig. 4).

The mRNA M:A encodes the classical membrane-bound IgE (Fig. 4). A donor site located 5 nt upstream of the stop codon for the secreted IgE (within the C4 exon) is ligated to the acceptor site of exon M1 and subsequently the M1 exon to the M2 exon (Fig. 4).

The mRNA M:B is spliced directly from the donor site of exon C4 to the acceptor site of exon M2. This results in an mRNA which lacks the entire M1 exon. The M2 exon is, through a frame shift generated by the splicing reaction from exon C4 to exon M2 (Fig. 4), translated in a different reading frame compared to the classical membrane transcript M:A. This results in a 136 aa-long carboxy-terminal extension of yet unknown biological function (Fig. 5A). The stop codon is located 326 nt downstream of the stop codon of M:A.

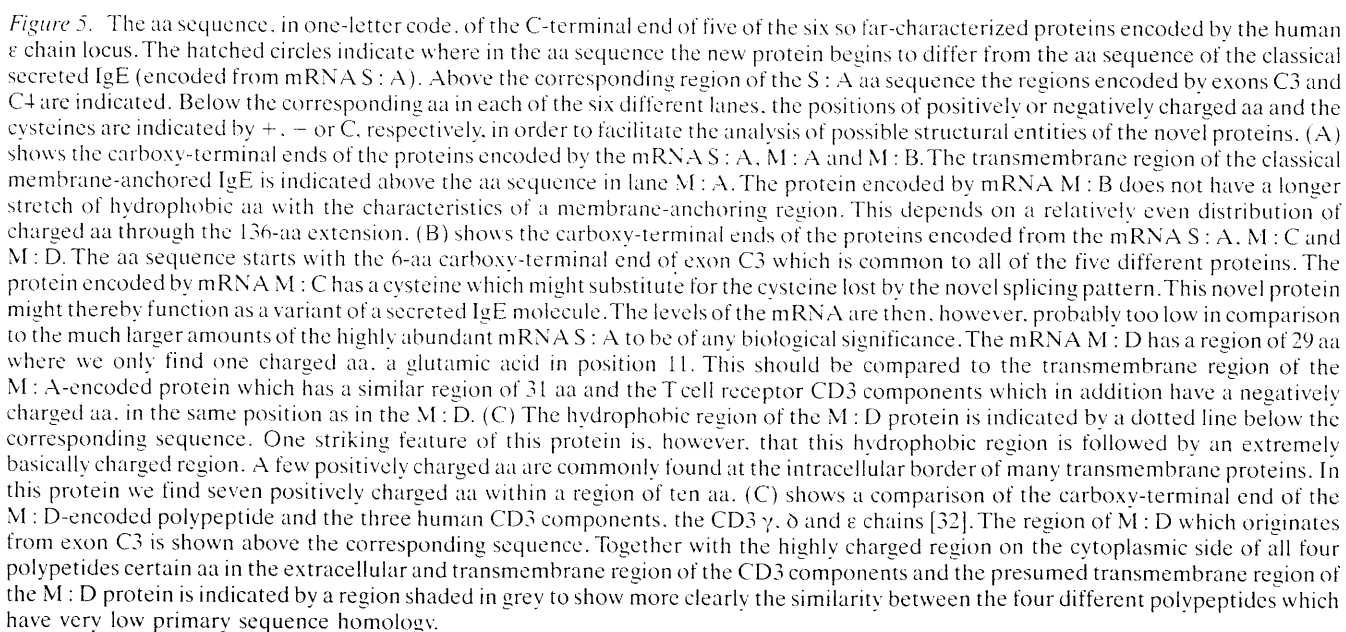
Analysis of mRNA M:C (Fig. 4) reveals two novel splice sites which are both located within the C4 exon. A new donor site is located 111 bp upstream of the stop codon for the secreted IgE and a new acceptor site 26 nt downstream of the same stop codon (Fig. 4). The open reading frame, from exon C4, continues through the new exon (the old 3' nc region) and terminates within the poly A signal of the C4 exon (the UAA of the AAUAAA) (Figs. 4 and 5).

The mRNA M:D (Fig. 4) lacks the entire coding region of exon C4. This occurs through a splicing reaction directly ligating the donor site of exon C3 to an acceptor site located in the 3' nc region of exon C4. This acceptor site is identical to the site used in the M:C mRNA (Fig. 4). The new exon, within the former 3' nc region of exon C4, is translated in a different reading frame compared to the M:C mRNA. The M:D mRNA does not utilize the classical acceptor site for the M1 exon but a novel acceptor site, which is located 156 bp upstream of the 5' end of this exon (Fig. 4). This new acceptor site is ligated to the donor site, located 8 nt downstream of the poly A addition site of S:A. The M:D mRNA is translated in a different reading frame through the M1 exon, compared to the M:A mRNA, and terminates 32 bp upstream of the donor site of this exon (Fig. 4 and 5B).

In U-266 we see a minority of the ϵ chain-specific mRNA in the region of the different membrane transcripts and the great majority in the region of approximately 2.4 kb where we find the transcripts for the secreted IgE (S : A: Fig. 3). This is in line with the finding that this cell line has been phenotypically characterized as an IgE-producing myeloma (plasma cell leukemia).

During B cell development the amounts of the separate mRNA change quite drastically from the early B cell, which mainly expresses the membrane-bound mRNA, to cells in the terminal plasma cell stage, a stage which expresses 95%–99% of its Ig mRNA in the secretory form. This shift in the relative amounts of the different mRNA is regulated at the level of differential splicing and possibly to some extent at the level of transcriptional termination.

Further dissection of the abundance of the different M2-containing transcripts shows that the lower band of the



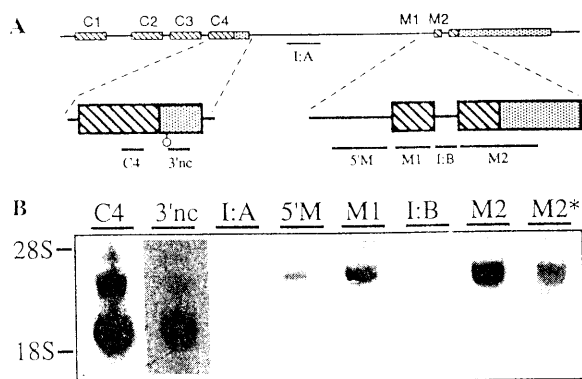


Figure 6. (A) shows a schematic map of the ϵ chain locus with the position of the different probes indicated below the corresponding region of the gene. (B) Poly(A)⁺ RNA was purified from the human IgE-producing B cell line U-266 and separated on a formaldehyde-containing agarose gel. Eight identical lanes were loaded on the same gel with 10 μ g of poly(A)⁺ RNA in each lane. The gel was blotted and the filter was cut into strips. These were hybridized separately with seven different probes originating from various regions of the human ϵ chain gene (see A). The position of the 18S and 28S ribosomal RNA is indicated as size markers. Two lanes marked M2 are present in the figure. The M2 lane shows the approximate amounts of the M2 hybridizing mRNA compared to the lanes 5'M and M1. A shorter exposure is included (lane M2*) to show more clearly the two different bands discussed in the results section.

two membrane transcripts is lacking the M1 exon (lane M1 in Fig. 6). Approximately 40% of the membrane transcripts are found in this region of the gel. The mRNA M:B (Fig. 4) probably represents the major transcript in this band. The oligonucleotide probe complementary to the 3' nc region hybridizes with a low but detectable signal to two or possibly three mRNA of different sizes in the region of the membrane-containing transcripts, in addition to a strong band for the mRNA S:A encoding the secreted IgE (Lane 3' nc in Fig. 6).

As can be seen from lane 5'M of Fig. 6 the region upstream of the M1 exon, which is present in the mRNA M:D, is expressed at quite significant levels.

The amount of the different mRNA has not been possible to measure with high accuracy except for the hybridizations done with probes directed against regions which are common to all transcripts. However, as we can easily detect most of these mRNA on Northern blot analysis (Fig. 6), we conclude that the mRNA are present at significant levels compared to the classical membrane transcript.

4 Discussion

The overall structure of the ϵ chain gene has changed only minimally during the 60–70 million years of evolutionary separation between rodents and man. The conservation of the coding regions, together with a high evolutionary drift of the intron sequences, is strongly indicative of an evolutionary advantage for the retention of the functional IgE mast cell system. The functional conservation, within a single species, of the high-affinity interaction between the IgE molecule and its receptor adds further support to this conclusion.

The existence of short regions of CA-containing repetitive sequences in the large C4/M1 intron indicates that in man there are remnants of the large blocks of CA repeats that are found in rodent C4/M1 introns. This implies that the CA repeats were present in this intron before the evolutionary separation of rodents and man. Does there exist a mechanism that keeps these repeats intact and nonmutated in the way we see in rodents? Does gene conversion contribute in maintenance of the integrity of certain repetitive sequences and do certain types of repetitive sequences actively get inserted into the genome? One example of a direct insertion of a CA repeat is found in the mouse Ig δ chain gene. This repeat is not present in the human counterpart and the CA repeat is flanked by 8-bp direct repeats [22].

The regions in close proximity to the CA repeats have, in rodents, gone through a large number of insertions and duplications. When we extend this comparison to include both rodents and man, the intron sequences have diverged to such an extent that no detectable homology can be found. This suggests that potential Z-DNA-forming sequences and possibly other types of repetitive sequences may stimulate rearrangements and thereby may speed up the evolutionary drift of regions in close proximity to these repeats. The existence of the large numbers of repetitive sequences found in all higher eucaryotes may thereby be of selective advantage. The low population sizes and the long generation times make the existence of mechanisms which speed up evolution a requirement for the development of higher eucaryotic species. The evolution of large gene families by gene duplication is one powerful mechanism in evolution. The existence of repetitive sequences, spread through the genome, may here increase the level of duplications by unequal "crossing-over" targeted at dispersed clusters of repetitive sequences.

We describe the characterization of four novel mRNA transcribed from the human IgE ϵ chain locus. One of the mRNA encodes a secreted IgE molecule identical to the classical secreted IgE but no biological function has yet been assigned to the proteins encoded by mRNA M:B, M:C and M:D. Recently one additional membrane-bound IgE was characterized by Peng et al. [14]. This novel splice variant, the M:E, encodes an IgE molecule which has 51 additional aa inserted between the C4 domain and the classical M1-encoded transmembrane region (Fig. 4). The authors show that this protein is present on the surface of the IgE-producing B cell. This molecule has transmembrane and cytoplasmic regions identical to the classical transmembrane IgE, indicating a functional identity.

What is the significance of the large number of mRNA transcribed from the human ϵ chain locus? By differential splicing the complexity of a locus can be greatly enhanced and a low level of presently erroneous splice combinations is probably of selectable advantage for the species as a source of potentially new proteins. Based on these ideas we favor the possibility that some of the variants are nonfunctional by-products of the complicated system of differential splicing.

However, for one of these novel transcripts our data suggest that its potential protein product might be functionally significant. The 3' nc region of exon C4 and the 156-bp 5' of exon M1 (both are present in the mRNA M:D and the latter in the mRNA M:E) are highly conserved between

rodents and man. The 3' nc region has a level of sequence conservation which is a few percent higher than any of the coding regions for the secreted IgE molecule (Fig. 3 B and [5, 23]). In addition the 156-bp region 5' to exon M1 has actually the same level of sequence conservation as the M1 exon itself (Fig. 3 B). When comparing regions just 3' or 5' to these novel exons we see that the level of sequence conservation falls to close to baseline levels (Fig. 3 B). The high degree of sequence conservation in the novel exons indicate that there are functional constraints to keep these regions conserved through evolution.

Data that on the other hand argue against a specific function is that even if all the splice sites described for these six different mRNA have been conserved between rat and man we see a relatively low level of conservation of the reading frames depending on a few frame shift mutations and a few stop codons in some of the reading frames. We can not, however, rule out the possibility that mutations could have affected the choice of splice sites and that we here have species-related differences.

The M:B, M:C and M:D transcripts are probably identical, in their 5' ends, to the mRNA for the classical secreted and membrane-bound IgE which is why we can, therefore, assume that they are efficiently translated. However, nothing is known concerning their stability and possible function. We find, by an analysis of the positions of charged aa and cysteines in the carboxy-terminal extensions of these proteins, a few interesting characteristics of these proteins (Fig. 5). The proteins encoded by M:B and M:C have no structure which resembles a transmembrane region indicating that, if they are stable after translation, they are secreted IgE molecules (Fig. 5). Another possibility, which at this time cannot be ruled out, is that they are unstable and degraded after translation and thereby represent nonfunctional products generated by errors in the differential splicing mechanism.

In the carboxy-terminal end of the protein encoded by M:D, we find a region which weakly resembles the transmembrane regions found in a group of T-cell-specific proteins, the CD3 γ , δ and ϵ chains (Fig. 5 B and C). The transmembrane region of M:D contains a negatively charged glutamic acid in position 11 from the outer membrane border which coincides with the position of a negatively charged aa found in all three CD3 components (Fig. 5 C). This negatively charged aa interacts with a positively charged lysine in the transmembrane region of the T-cell receptor α and β chains thereby stabilizing the interaction between the different T-cell receptor components. The potential transmembrane region is preceded by a cysteine and a few charged aa in similar positions as for the CD3 components, and is followed by a strongly positively charged region with seven arginines (Fig. 5 C). The latter is characteristic of cytoplasmic borders of many transmembrane proteins. At this stage, we can only speculate on the presence of a novel signal-transducing molecule which interacts with the presumptive protein product of the mRNA M:D.

Recently the nt and aa sequences for the membrane IgM were published for two different species of bony fish (teleosts) [24, 25]. In these two species the first membrane exon is spliced directly to the C3 exon, thereby deleting exon C4 in a manner similar to that is seen in the ϵ transcript M:D. However, the presumptive transmembrane region of M:D has no direct aa sequence homology to the classical membrane-anchoring tail of the different surface Ig. This indicates that we have a potentially new signal-transducing system for IgE-producing B cells which clearly differs from the classical membrane-bound Ig. Other possibilities are that the M:D is a remnant of an ancestral membrane Ig-encoding mRNA or, as we at present cannot rule out, a nonfunctional protein product of the differential splicing mechanism. Fusion proteins are being prepared for production of specific antisera for each of the new proteins in order

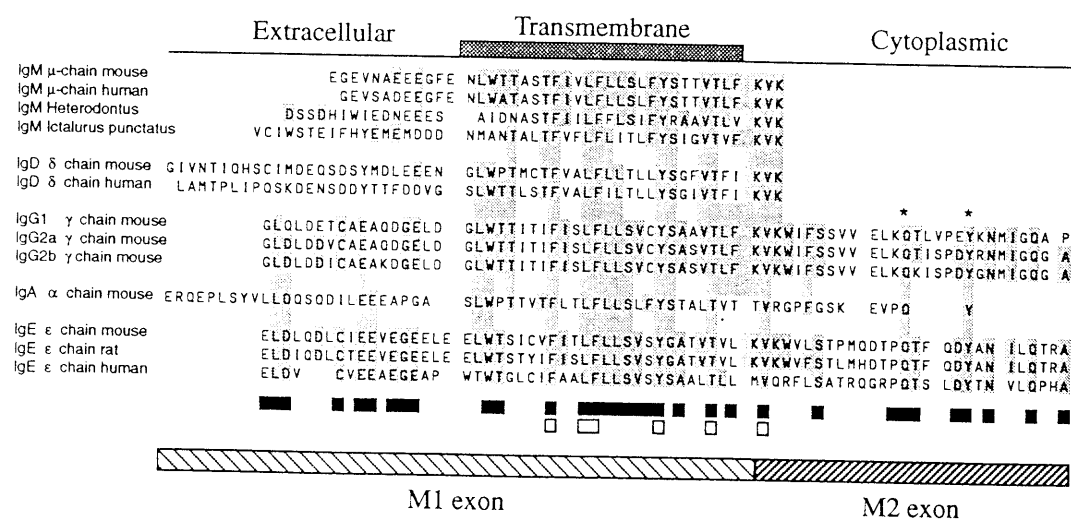


Figure 7. The figure shows a comparison of a panel of Ig membrane domains from fish to higher mammals. All aa regions encoded by the M1 and M2 exons are depicted in the figure. The aa regions showing a high degree of conservation are shown within grey boxes. aa conserved between mouse, rat and human ϵ chains are shown by filled boxes below the corresponding sequence in the lower part of the figure. All aa which are conserved between the transmembrane regions are shown by open boxes below the corresponding sequence. The conserved glutamine and tyrosine in the cytoplasmic region of all non- μ , non- δ Ig classes is indicated by an asterisk above the corresponding aa. The aa encoded by the separate exons is shown in the bottom of the figure. This is valid for all Ig classes except for the α chain where the complete region is encoded by one single exon (M).

to answer questions concerning their location, stability and possible function.

In Fig. 7 the aa sequence of the transmembrane and cytoplasmic regions of human IgE is compared to a panel of membrane Ig characterized in mouse, man and two species of fish. An exceptionally high degree of sequence conservation is found in the membrane-spanning region of the antibodies from fish to man (Fig. 7). This part of the molecule has been found to be crucial for the interaction with accessory signal-transducing molecules [26–28].

All antibodies formed after a switching event (non- μ and non- δ) have an intracellular extension of 14 (α chain) or 28 aa (γ and ϵ chains) attached to their carboxy terminus [13]. This can be compared to only three aa for the μ and δ chains [13]. Nothing is yet known about the function of this cytoplasmic extension. The α chain is relatively distantly related to the γ and ϵ chains, which are highly homologous to each other (Fig. 7). This is also reflected in the organization of the exons where the α chain membrane region is encoded by one exon compared to two exons for both the γ and ϵ chains. A few interesting characteristics are found upon a more detailed analysis of these peptide extensions. We find one glutamine and one tyrosine residue conserved between all non- μ and non- δ cytoplasmic membrane extensions (Fig. 7). Tyrosine residues have been implicated to play a role in the internalization of membrane-bound receptor molecules. One example is found in the low-density lipoprotein receptor where a cytoplasmic tyrosine has been shown to be vital for rapid endocytosis [29].

The coupling of signal-transducing molecules to the membrane-bound forms of the Ig may be influenced by a difference in the structure of the cytoplasmic tail of the antibody. Three membrane proteins have been found to interact with the surface-bound IgM molecule and they are all phosphoproteins. In at least one of these three polypeptides the phosphate is bound to a cytoplasmic tyrosine [27]. There is, however, no evidence that this phosphate has a regulatory function.

Some of the different signal-transducing molecules differ in size between the μ and the δ chains [27, 30]. Recent data from Venkitaraman et al. [31] show that all Ig isotypes are associated with proteins encoded by the genes cloned as the B cell-associated proteins mb-1 and B-29 and that the difference in size of the α chains (mb-1) associated with μ and δ chains are due only to a difference in glycosylation. They show that mb-1 and B-29 are sufficient for surface expression of the membrane Ig of all isotypes. Whether these are the only molecules associated with all the different surface Ig is not yet clear. The cytoplasmic region of the membrane Ig may substitute for functions performed by as yet uncharacterized accessory molecules for the μ and δ chains or may perform additional functions needed for B cells which have gone through an isotype switching event.

I would like to thank Prof. Jan Andersson for support during the project and Prof. Hans Bennich and Dr. Stanley Burnett for revision of the manuscript. I would like to thank Maria Aveskog for skillful technical assistance.

Received February 14, 1992; in final revised form July 15, 1992.

5 References

- 1 Johansson, S. G. O. and Bennich, H., *Immunology* 1967, 13: 381.
- 2 Ishizaka, K. and Ishizaka, T., *J. Immunol.* 1967, 99: 1187.
- 3 Hagan, P., Blumenthal, U. J., Dunn, D., Simpson, A. J. G. and Wilkins, H. A., *Nature* 1991, 349: 243.
- 4 Ishida, N., Ueda, S., Hayashida, H., Miyata, T. and Honjo, T., *EMBO J.* 1982, 1: 1117.
- 5 Steen, M.-L., Hellman, L. and Pettersson, U., *J. Mol. Biol.* 1984, 177: 19.
- 6 Flanagan, J. G. and Rabbitts, T. H., *EMBO J.* 1982, 1: 655.
- 7 Max, E. E., Battey, J., Ney, R., Kirsch, I. R. and Leder, P., *Cell* 1982, 29: 691.
- 8 Sun, L. K., Liou, R. S., Gossett, L. A., Sun, C., Davies, F. M., MacGlashan, D. W. and Chang, T. W., *J. Immunol.* 1991, 146: 199.
- 9 Hellman, L., Steen, M.-L., Sundvall, M. and Pettersson, U., *Gene* 1988, 68: 93.
- 10 Islam, K. B., Nilsson, L., Sideras, P., Hammarström, L. and Smith, C. I. E., *Int. Immunol.* 1991, 3: 1099.
- 11 Nilsson, L., Islam, K. B., Olafsson, O., Zalcberg, L., Samakovlis, C., Hammarström, L., Smith, C. I. E. and Sideras, P., *Int. Immunol.* 1991, 3: 1107.
- 12 Rothman, P., Li, S. C., Gorham, B., Glimcher, L., Alt, F. and Boothby, M., *Mol. Cell. Biol.* 1991, 11: 5551.
- 13 Word, C. J., Mushinski, J. F. and Tucker, P. W., *EMBO J.* 1983, 2: 887.
- 14 Peng, C., Davies, F. M., Sun, L. K., Liou, R. S. and Chang, T. W., *J. Immunol.* 1992, 148: 129.
- 15 Flanagan, J. G. and Rabbitts, T. H., *Nature* 1982, 300: 709.
- 16 Sanger, F., Niclen, S. and Coulson, A. R., *Proc. Natl. Acad. Sci. USA* 1977, 74: 5463.
- 17 Auffray, C. and Rougeon, F., *Eur. J. Biochem.* 1980, 107: 303.
- 18 Maniatis, T., Fritsch, E. F. and Sambrook, J., in *Molecular Cloning, A laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1982.
- 19 Brenner, C. A., Tam, A. W., Nelson, P. A., Engleman, E. G., Suzuki, N., Fry, K. E. and Larrick, J. W., *BioTechniques* 1989, 7: 1096.
- 20 Nilsson, K., Bennich, H., Johansson, S. G. O. and Ponten, J., *Clin. Exp. Immunol.* 1970, 7: 477.
- 21 Hellman, L., Josephson, S., Jernberg, H., Nilsson, K. and Pettersson, U., *Eur. J. Immunol.* 1988, 18: 905.
- 22 Word, C. J., White, M. B., Kuziel, W. A., Shen, A. L., Blattner, F. R. and Tucker, P. W., *Int. Immunol.* 1989, 1: 296.
- 23 Hellman, L., Pettersson, L., Engström, Å., Karlsson, T. and Bennich, H., *Nucleic Acids Res.* 1982, 10: 6041.
- 24 Wilson, M. R., Marcuz, A., van Ginkel, F., Miller, N. W., Clem, L. W., Middleton, D. and Warr, G. W., *Nucleic Acids Res.* 1990, 18: 5227.
- 25 Bengtén, E., Leandersson, T. and Pilström, L., *Eur. J. Immunol.* 1991, 21: 3027.
- 26 Shaw, A. C., Mitchell, R. N., Weaver, Y. K., Campos-Torres, J., Abbas, A. K. and Leder, P., *Cell* 1990, 63: 381.
- 27 Campbell, K. S. and Cambier, J. C., *EMBO J.* 1990, 9: 441.
- 28 Hombach, J., Leclercq, L., Radbruch, A., Rajewsky, K. and Reth, M., *EMBO J.* 1988, 7: 3451.
- 29 Davies, C. G., van Driel, I. R., Russell, D. W., Brown, M. S. and Goldstein, J. L., *J. Biol. Chem.* 1987, 262: 4075.
- 30 Wienands, J., Hombach, J., Radbruch, A., Riesterer, C. and Reth, M., *EMBO J.* 1990, 9: 449.
- 31 Venkitaraman, A. R., Williams, G. T., Dariavach, P. and Neuberger, M. S., *Nature* 1991, 352: 777.
- 32 Clevers, H., Alarcon, B., Wileman, T. and Terhorst, C., *Annu. Rev. Immunol.* 1988, 6: 629.