Chemica Scripta 1986, **26B**, 343–349

Evolutionary Aspects of Immunoglobulin-related Genes

Susumu Tonegawa and Haruo Saito*

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Paper presented by Susumu Tonegawa at the Conference on 'Molecular Evolution of Life', Lindingö, Sweden, 8–12 September 1985

Abstract

In the vertebrates' immune system the recognition of nonself molecules is carried out by two sets of evolutionarily related glycoproteins, immunoglobulin (Ig) and T cell receptors (TCR). The most characteristic feature of these proteins is their enormous variability. Unlike most other known genes, an Ig or TCR gene is carried in the germline as gene segments (*V* and *J* or *V*, *D* and *J*) which are assembled at the DNA level specifically during lymphocyte differentiation. This somatic recombination generates much of the variability observed in the so called variable (V) regions of these polypeptide chains.

Ig and TCR polypeptide chains are composed of internally repeated common homology units which fold into similar globular domains: The V regions of all these chains, and the C (constant) regions of Ig light and TCR chains consist of a singled homology unit, while the C regions of Ig heavy chains contain three or four homology units. At the DNA level each of the C region homology units corresponds to an independent exon. As to the V homology units, the coding DNA is split into two (*V* and *J*) or three (*V*, *D*, and *J*) segments in the germline but composes a single exon in the B and T lymphocytes, for Ig and TCR genes, respectively. These results suggest that: (1) both Ig and TCR genes evolved from a common primordial gene of the size of one homology unit as a result of a series of duplication, sequence divergency, and fusion of the duplicated DNA copies by RNA splicing; (2) introns of some eukaryotic genes arose from the flanking non-coding sequences of ancient genes subsequent to gene duplication; (3) *V–J* (and possibly *V–D–J*) joining is a reversal of an ancient insertion of a transposition-like DNA element into one of the duplicated primordial genes.

Ig and TCR genes are two of an increasingly larger number of evolutionarily related genes, most of which encode integral membrane proteins, although somatic recombination seems to be restricted to the two gene families. These findings suggest that the Ig-like combining sites and/or Ig-like domains are widely used by cell surface proteins either for ligand binding or cell-cell interaction. The finding also suggests that many eukaryotic proteins may be classified into evolutionarily related but functionally distinct families.

Introduction

Vertebrates are endowed with a highly effective body defense mechanism referred to as the immune system. The critical event in mounting the immune response is the recognition of chemical markers present on the surface of molecules or a collection of molecules that are foreign to the body (antigens). This task is entrusted to two groups of proteins, namely immunoglobulins synthesized and secreted by B lymphocytes and the so called T cell receptors present on the surface of T lymphocytes. The most intriguing properities of these proteins is their variability of structure: A single organism can synthesize tens and hundreds of millions of structurally different antibodies and T cell receptors. During the past ten years the analysis of genes coding for these proteins uncovered not only the genetic origins of the enormous variability but also several intriguing features of the behaviour of these genes in evolution. Some of these findings are highly relevant to our understanding of the evolution of eukaryotic genes in general.

Structure of antibody

The basic unit of an immunoglobulin molecule is a tetramer composed of two identical light chains and two identical heavy chains. Both chains consist of linear arrays of a homology unit of about 110 amino acid residues: a light chain is a dimer while a heavy chain is a tetramer or a pentamer. The homology units located at the amino terminal ends of both chains are diverse in the sequence (variable or V region) and are primarily responsible for the determination of antigen-binding specifications. The vast majority of antibody variability is in the V regions. The other homology units form the constant (C) regions which define the 'type' and 'class' of immunoglobulin molecules and carry polypeptides involved in specific effector functions of the immunoglobulin molecule. A given mammalian antibody molecule contains one of two types of light chains, κ, or λ, and one of five classes of heavy chains, μ, γ, α, etc. At the three dimensional level each homology unit folds into a relatively independent globular domain characterised by several antiparallel β strands. The domain is stabilized by a disulfide linkage provided by a pair of cysteins universally present in the immunoglobulin homology units. The antibody combining site can be viewed as a cleft formed between the light and heavy chain V domains that are held together by characteristic noncovalent bonds. Realization of the occurrence of internally repeated units in the polypeptide chains lead to the hypothesis that the immunoglobulin genes arose by a series of duplications followed by sequence divergence of a promodial gene corresponding in size to a single homology unit. As seen below this idea was born out by the direct structural analysis of the genes.

Structure of antibody genes and the origin of introns

Mammalian immunoglobulin chains are encoded in three unlinked gene families: λ light-chain genes, κ light-chain genes, and heavy-chain genes residing on separate chromosomes. The most intriguing feature of these genes is their dynamic behaviour in the life-span of an individual organism. In fact none of these genes is carried in the germ line in the

* Present address: Division of Tumor Immunology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02115, USA.

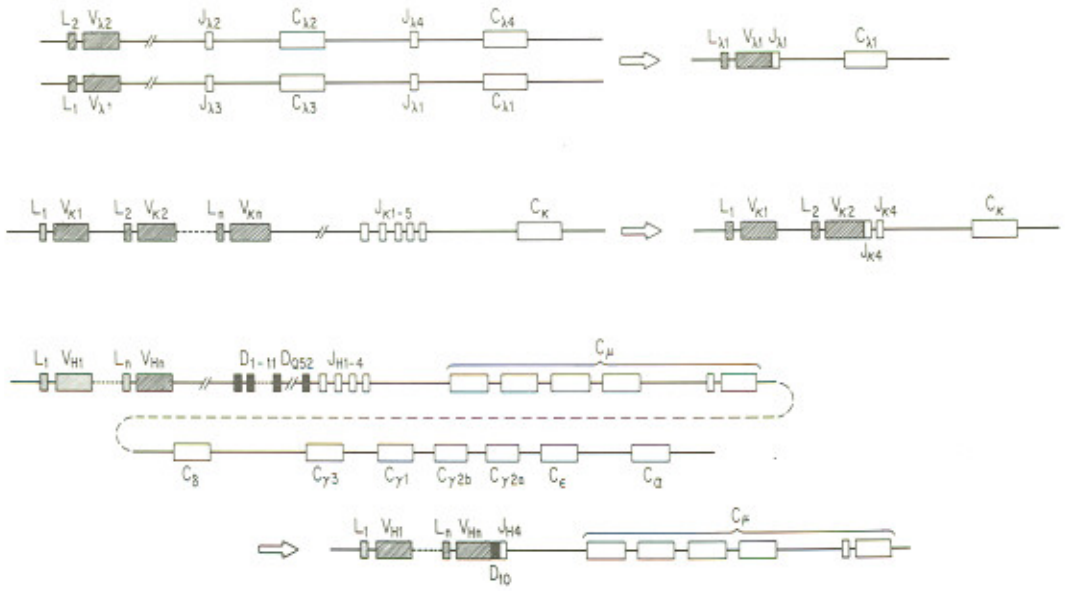


Fig. 1. Organization and somatic reorganization of immunoglobulin gene families. The diagram is for mouse Balb/c. The exon structures of heavy chain genes other than μ is abbreviated. Arrows indicate the change from germline configuration to the configuration observed in B lymphocytes.

complete form. Rather, they are transmitted through generations as gene segments which undergo a series of highly characteristic recombinational events which are strictly restricted to the differentiation of B lineage cells. More specifically, as shown in Fig. 1, the coding potential of the V regions of a light and heavy chains are split in the germline into two (*V* and *J*) and three (*V*, *D* and *J*) gene segments, respectively [1]. In B lineage cells which produce immunoglobulin chains, the corresponding gene segments are assembled into a continuous stretch of DNA as a consequence of *V–J* or *V–D–J* joining. Since a given immunoglobulin gene family contains multiple and different copies of each of the two (*V* and *J*) or three (*V*, *D* and *J*) types of gene segments and that *V–J* or *V–D–J* joinings can occur in a variety of combinations, a large number of different V regions can be generated starting with a relatively limited number of gene segments carried in the germ line. Enzymology of the recombinational events is poorly known. However, it is highly likely that B lymphoid cells contain a specific recombinase that recognizes specific sequences in or around the recombination sites. Indeed the nucleotide sequencing studies revealed that a characteristic heptamer and nonamer sequences are conserved adjacent to the recombination sites in the non-coding regions.

Another important finding made during the analysis of immunoglobulin genes is concerned with the positions of introns. It was shown that each homology unit is encoded by an independent exon [2]. Furthermore the signal peptide present at the amino terminal end of a nascent immunoglobulin chain and the so called hinge region of immunoglobulin γ chains are encoded by their own exons. On the basis of these findings we postulated that some eucaryotic genes may have been generated by shuffling of exons each of which encodes a polypeptide with a distinct function and/or a relatively independent three dimensional structure [3]. According to this hypothesis the origin of introns is non protein-coding sequences that flank mobile exons (see below). Interestingly such shuffling of exons at the DNA level does occur during the late stage of B cell development within the immunoglobulin

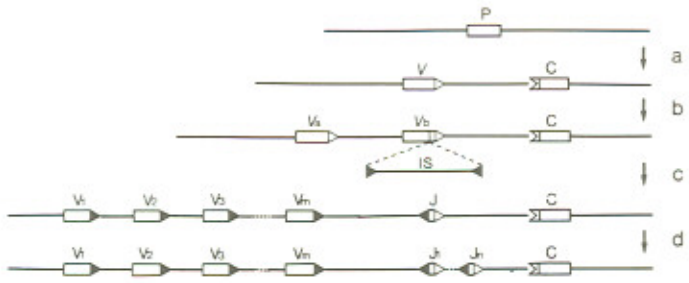


Fig. 2. Evolution of immunoglobulin and T cell receptor genes. 'IS' designates a transposition-like element thought to be inserted to one of the multiple DNA copies for a V region.

heavy chain gene family leading to a 'switch' of the expression of the heavy chain from one class (such as μ) to another (such as γ). While not all introns may have arisen in evolution by the proposed scheme, some recent molecular genetic studies on proteins with internal repetitions support the idea that some introns indeed are generated by gene duplications [4–6].

V–J joining may be a reversal of an ancient insertion event

It is reasonable to assume that the light chain gene containing two homology units was generated by duplication of a primordial gene of the size of a single homology unit followed by divergency and fusion of the two units. Let me consider these processes somewhat more in detail. Since the major source of gene duplication is an unequal crossing over between two homologous stretches of DNA it is very unlikely that the primordial gene corresponds exactly to the duplication unit. Rather, the flanking sequences are also duplicated, leaving a spacer between the two copies of the duplicated genes. Sakano *et al*, assumes that a 'dimeric gene' coding for two immunoglobulin domains was created from the duplicated primordial gene by exploitation of a pre-existing RNA splicing mechanism [7]. (Fig. 2) In this ancient *V–C* gene the *J* gene segment is thought to be an integral part of the ancestral *V* DNA segment because the J region is homologous to the caboxyl end of the C region. The ancestral *V* DNA,

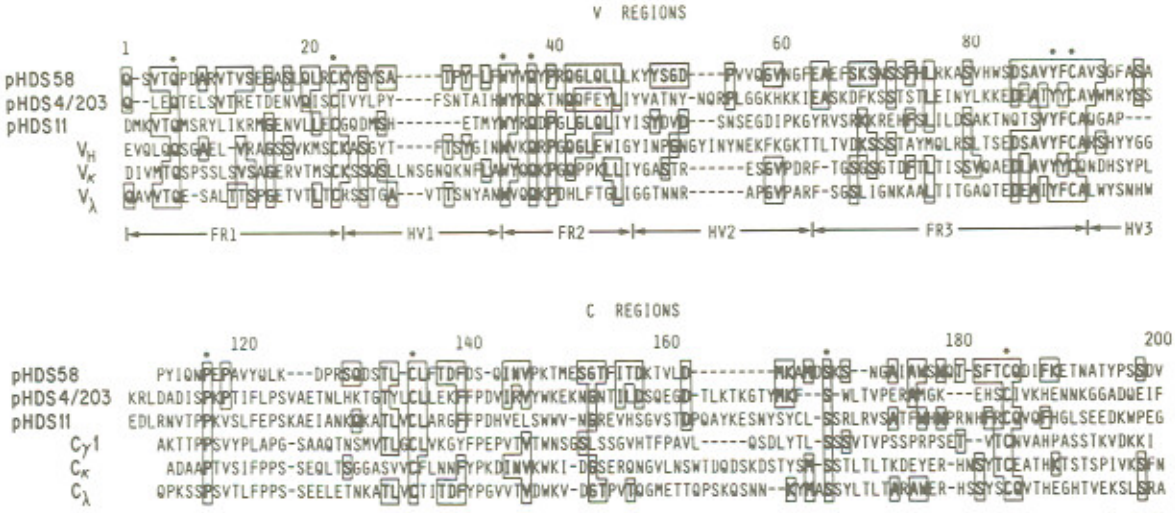


Fig. 3. Amino acid sequence homology among T cell receptor α (pHDS58), β (pHDS11), and γ (pHDS4/203) chains and immunoglobulin chains. Stars designate the residues conserved universally.

together with some flanking sequences, duplicated, triplicated, and so on under pressure to increase V-region diversity. As splicing occurs most efficiently as an intramolecular reaction, one requirement of such a 'polymeric gene' would be that the *V* and *C* DNA sequences are co-transcribed. In the long RNA transcript every V-coding sequence would have a splice donor site at its 3′ end, so that a series of mature mRNA containing different V-coding sequences and the same C-coding sequence could be generated. However this strategy of increasing variability seems to have severe limitations by at least two reasons. First, as the number of *V* DNA segments increases the transcription unit would become intolerably long. Second, as the gene grows it would become increasing difficult to restrict the expression to only one *V* gene segment, which would be required to fulfill the monospecificity of B lymphocytes.

Sakano *et al*, believe that these limitations were resolved by the invention of the somatic *V–J* joining event [7] (Fig. 2.) They propose that such a mechanism was initiated when an IS-like DNA element was accidentally inserted into one of the multiple *V* DNA copies of the ancestral polymeric gene, splitting it into two portions, one corresponding to the present day germ line *V* DNA segment and the other to the *J* DNA segment. While this insertion was fixed in the germ line genome, a mechanism to excise the inserted DNA and rejoin the split V-coding sequences was established in lymphocytes. As excision is a common and major manifestation of prokaryotic IS elements, the basic mechanism for such a process would probably have been available in vertebrates. Once an ancestral *V* DNA is split, the major body of the V-coding sequence and the 3′ flanking sequence containing one end of the IS-like DNA element can multiply independently of the *J* DNA segment. Becaused every copy of the duplicated *V* DNA segment would have the excision-recombination site at its 3′ end, it can rejoin with the *J* DNA segment which is kept in the vicinity of the C-coding DNA. Because the rearranged DNA can bring in its own transcription promoter, it is no longer necessary to confine the duplicated *V* DNA segments and the *C* DNA segment within the same transcription unit in the germ line genome. Furthermore this genetic strategy can provide a means to restrict the expression of immunoglobulin chains to that encoded by the rearranged *V* gene segment. RNA transcribed from unrearranged *V* DNA segments will not be spliced to that from the *C* DNA segment because the former RNA lacks the RNA splice signal.

T cell receptor genes

Unlike immunoglobulins which recognize and bind free antigens and are responsible for humoral immunity, T cell receptors recognize cell-bound antigens in the specific molecular context of self major histocompatibility complex (MHC) products and are responsible for cellular immunity [8–10]. The MHC-restriction appears to be largely acquired by a differentiating T cell population under the influence of MHC antigens expressed in the thymus, suggesting that precursor T cells are selected on the basis of their reactivity with MHC determinants expressed in the host thymus (thymic education) [11–13]. In order to understand the molecular basis of MHC restriction and thymic education it is of paramount importance to determine the structure of a T cell receptor. For many years T cell receptors were very elusive. However the first glimpse of these molecules was obtained in 1982 and 1983 through the development of effective antibodies [14–16]. These studies lead to the conclusion that the portion of the receptor determining specificity is a heterodimeric glycoprotein corresponding to a molecular weight (M_r) of about 90 K consisting of a 40–45 K α subunit and a 40–45 K β subunit. Moreover, peptide fingerprint analysis suggested that both the α and β subunits, like immunoglobulin heavy and light chains, are composed of variable (V) and constant (C) regions.

More recently the genes coding for these polypeptide chains were cloned [17–20]. The nucleotide sequences of these DNA clones permitted the deduction of the complete primary structure of the α and β subunits. When the predicted amino acid sequences of the α and β chains were compared with all other known protein sequences it was discovered that these chains are significantly homologous to immunoglobulin chains, particularly light chains (30–35% homology) (Fig. 3). The conserved residues include the two pairs of cysteines that are involved in the intradomain disulfide linkages in the V and C regions, the invariant aromatic rings Trp (34), Tyr (88), and

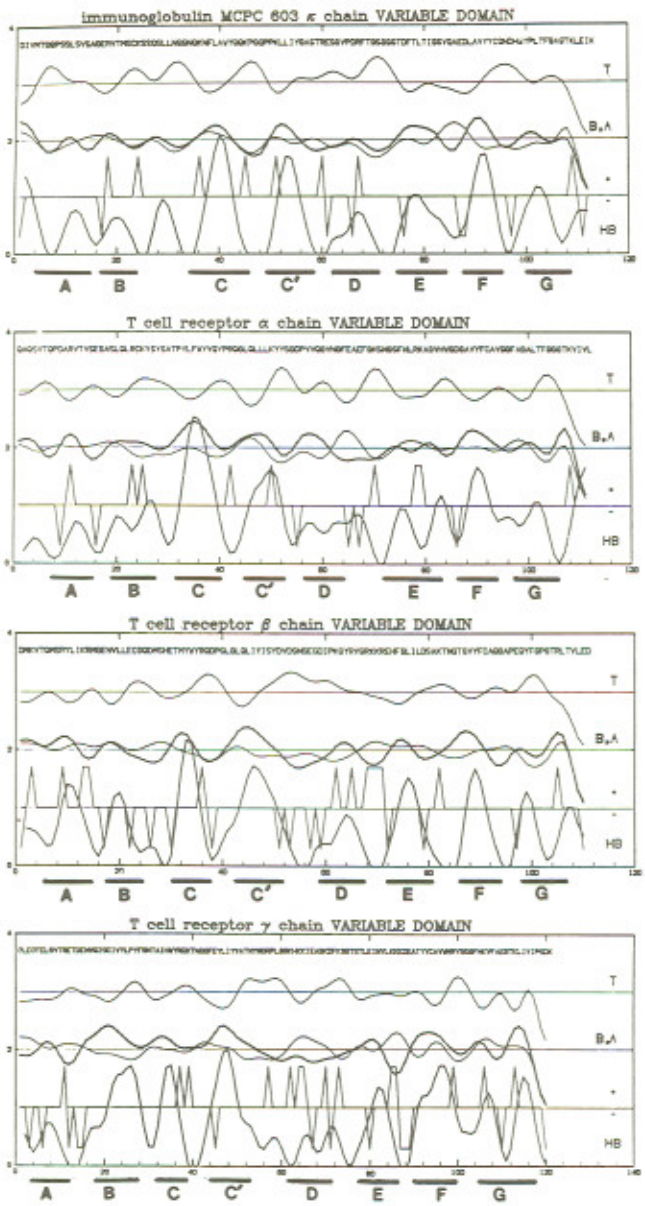


Fig. 4. 'Secondary Structure Profiles' of immunoglobulin and T cell receptor V regions. Each box shows from top to bottom, reverse turn propensity, T; α-helix (thin line A) and β sheet (heavy line B) propensities; profile of charged residues (+ −); and hydrophobicity. These were computed by J. Novotný of Massachusetts General Hospital in Boston as described [42]. Bars below the boxes indicate β strands identified from X-ray data (MOPC 603 V_L domain [43]) or deduced from the profiles (T cell receptor chains).

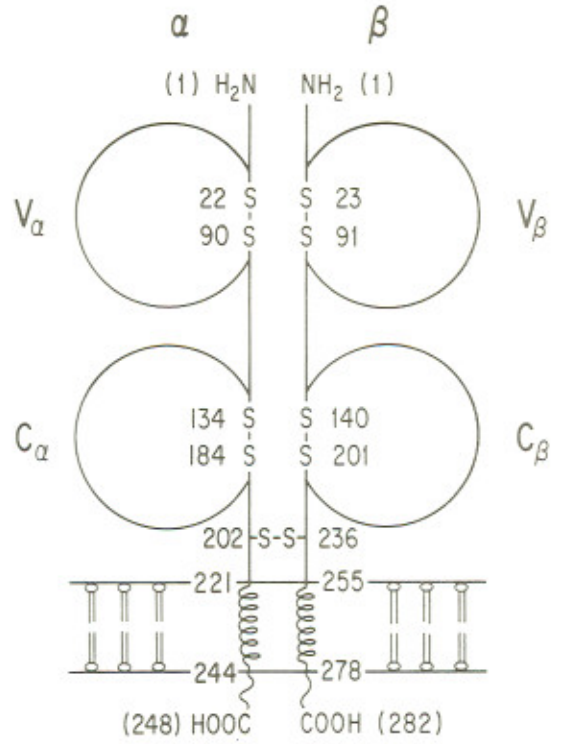


Fig. 5. Structure of the T cell receptor αβ heterodimer on the surface of a cloned functional CTL, 2C. Numbers represent the amino acid positions. 'S' represents the sulfur involved in intrachain or interchain disulfide linkages.

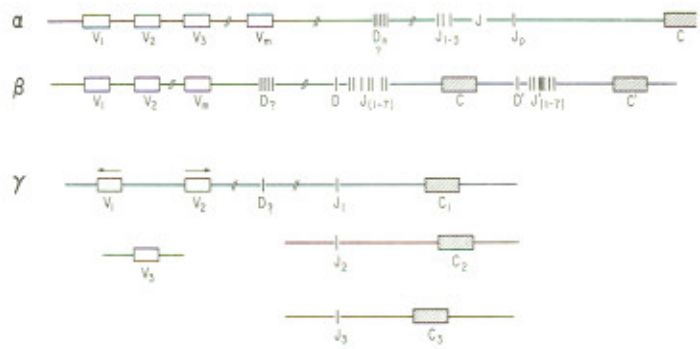


Fig. 6. Germline organization of T cell receptor genes α, β and γ.

Phe (101) at V–V interface, and the invariant Gln (37) that form a hydrogen-bond across the V–V interface. Furthermore the 'secondary structure profile' of the V_α and V_β regions (see Fig. 4) strongly suggest that these V regions are folded into the typical immunoglobulin domains characterized by anti-parallel β strands [21]. Figure 5 depicts the gross structure of the αβ heterodimer deduced from these analyses.

The striking structural similarity between immunoglobulin chains and T cell receptor subunits parallels with the equally striking resemblance of the organization and structure of the corresponding genes. As summarized in Fig. 6 the α and β chain genes are also split to *V*, (*D*), *J*, and *C* gene segments in the germline genome [22–25]. In T cells a selected set of a *V*, (*D*), *J* gene segments are assembled to form a continuous stretch of DNA coding for the V region. Furthermore the T cell receptor gene segments seem to share with immunoglobulin gene segments the same set of conserved heptamer and nonamer sequences as the signals for the V–J or V–D–J joining, although some interesting variations can be seen in the association of the spacer lengths and the gene segments.

In conclusion, immunoglobulins the T cell receptors show clear resemblance both at the level of the protein and DNA. It is inescapable to conclude that they have diverged from a common primordial gene.

T cell γ gene

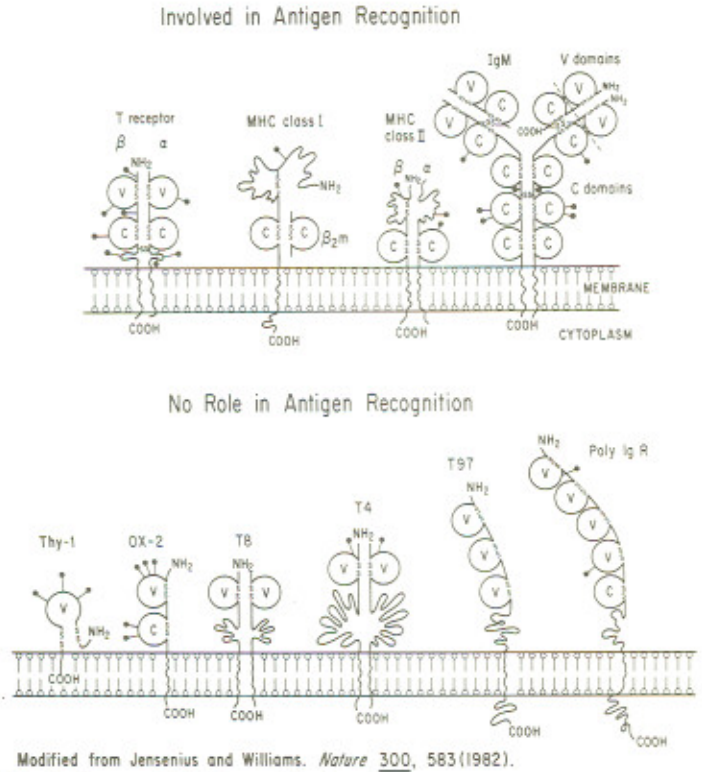
During the search for the α and β chain gene we encountered the third gene (referred to as γ) expressed in the same cytotoxic T cell clone, 2C from which the α and β cDNA clones were isolated. [26] This gene has many properties in common with α and β genes which include [26–28]: (1) assembly from gene segments resembling the gene segments for immunoglobulin V, J, and C regions; (2) rearrangement and expression of these genes segments in T cells and not in B cells; (3) low but distinct sequence homology to immuno-

Table I. *Cell surface proteins with immunoglobulin-like domains*

	Polypeptide chains		Function	Distribution	DNA rearrangement
1	Immunoglobulin	κ light	Antigen recognition	B lymphocytes	
2		λ light			+
3		Heavy			
4	T cell receptor	α	Antigen and MHC	T lymphocytes	
5		β	recognition		+
6		γ			
7	MHC Class I	Heavy	Restriction of	Many cell types	
8		β_2M	T cell response	but low in neuronal or glial cells	−
9	MHC Class II	α	Restriction of both	B cells, activated	
10		β	B and T cell responses	macrophages, dendritic cells, ephitherial cells	
11	T4 (L3T4)		Unknown	Most T helper cells	−
12	T8 (Ly2)		Unknown	Most T killer and T suppressor cells	−
13	Thy 1		Unknown	Thymocytes, neuronal cells and others	−
14	OX-2		Unknown	Thymocytes, brain cells and others	−
15	Poly Ig receptor		Transcellular transport of poly IgA and IgM	Glandular epithelial cells	−

globulin V, J, and C regions; (4) conservations of key residues involved in domain folding and interdomain interaction; (5) secondary structure profile; (6) sequences reminiscent of the transmembrane and intracytoplasmic regions of integral membrane proteins; (7) a cystein residue at the position expected for a disulfide bond linking two subunits of a dimeric membrane protein molecule. In spite of the long list of similarity between the γ gene and the α and β genes there seems to be some features unique to the γ gene which include [29–32]: (1) the diversity of the germline gene segments is significantly limited (see Fig. 6); (2) the expression seems to be dispensable for some helper T cells; (3) the expression is primarily in immature T cells while the β gene expression is constitutive throughout T cell differentiation and the α gene expression is primarily in mature T cells.

While no definitive function has been assigned to the putative γ chain we believe that the available data are best interpreted by assuming that the γ chain constitutes at least part of the receptor on immature T cells that undergo intrathymic selection (i.e. thymic education). Regardless whether this idea is correct it is intriguing that this gene and its putative gene product is so analogous to the α and β genes and their gene products.



Modified from Jensenius and Williams. *Nature* 300, 583 (1982).

Fig. 7. Immunoglobulin superfamily.

Other immunoglobulin-like proteins and genes

Prior to the discovery of the T cell receptor α, β, and γ chains, several proteins were shown to have significant homology to immunoglobulin chains. These include the Class I and Class II MHC proteins [33, 34], the T cell surface protein Thy-1 [35], and the poly IgA and IgM receptor which is present on the surface of glandular epitherial cells and mediates transcellular transport of these immunoglobulins (Table I) [36]. Furthermore more recent studies are expanding the list of immunoglobulin-like proteins. For instance the T4 and T8 antigens present on the surface of most human helper T cells and cytotoxic (and suppressor) T cells, respectively, each contains

Table II. *Similarities between Ig V-like domains*

V_L	V_A	OX-2	T8	Thy-1	IgRc	T97-1	T97-2	T97-3	
31	33	30	25	23	22	20	22	24	V_H
	34	29	33	24	21	19	22	24	V_L
		27	28	22	31	21	23	20	V_A
			22	29	21	22	25	15	OX-2
				25	26	25	28	27	T8
					22	22	23	19	Thy-1
						22	14	17	T97-1
							16	23	T97-2
								18	T97-3

V_H, Human NEWM; V_L, mouse lambda, MOPC104E; V_A, mouse T cell receptor alpha, HDS58; Thy I, Rat Thy-1; IgRc, poly Ig receptor (4th domain).

Fig. 8. Polypeptide chains with VJ region but without gene rearrangement.

an immunoglobulin V-like domain [37, 38, 39]. (Table I and Fig. 7). In addition OX-2 glycoprotein known to be present on the surface of rat thymocytes and some brain cells is composed to one V- and one C-like domain plus a transmembrane and intracytoplasmic peptides [40] (Table I and Fig. 7). Furthermore we have recently identified a cDNA clone (T97) in the library prepared from the cytotoxic T cell clone, 2C. The T97 gene seems to be expressed specifically in T cells and its putative gene product contains three immunoglobulin V-like domains, a long (about 160 residues) non immunoglobulin-like sequence, a transmembrane peptide and an intracytoplasmic peptide (Fig. 7).

The sequence homology among these immunoglobulin-like polypeptide chains can be as low as 14% (T97 domain 2 *vs* T97 domain 1) but in most cases it falls to a value between 20 and 35% (see Table II). In those cases where the sequence homology to an immunoglobulin domain is relatively low some of the residues central to the immunoglobulin fold such as the invariant cysteins and tryptophane are missing (example, the N-terminal homology units of the Class II MHC products and Thy-1, etc). In these cases it is highly unlikely that the homology unit in question will fold into the immunoglobulin-like domain. On the other hand it is virtually certain that some other homology units such as those at the N-terminal of T4 and T8 fold into immunoglobulin-like domains (Figs. 4, 8).

Conclusions

A great deal has been learned by molecular genetic studies of immunoglobulins and T cell receptors. The genetic origins of antibodies, one of the most debated problems in immunology has now been resolved at least in the outline: somatic recombination among germline-carried gene segments and somatic hypermutation (not discussed in this article, see ref. 1) in the assembled immunoglobulin genes play key roles in the generation of the tremendous diversity seen in the population of antibodies synthesized by a single organism. Another problem, also heatedly debated in immunology, namely whether a T cell uses an immunoglobulin as its antigen receptor is now resolved beyond doubt: It does not. Instead, T cell receptors are encoded by two sets of genes that are clearly distinct from immunoglobulin genes but share many features in common with them including somatic rearrangement. At least one additional set of genes (γ), also immunoglobulin gene-like, may also be involved in the determination of T cell specificity in some stages of T cell differentiation. The

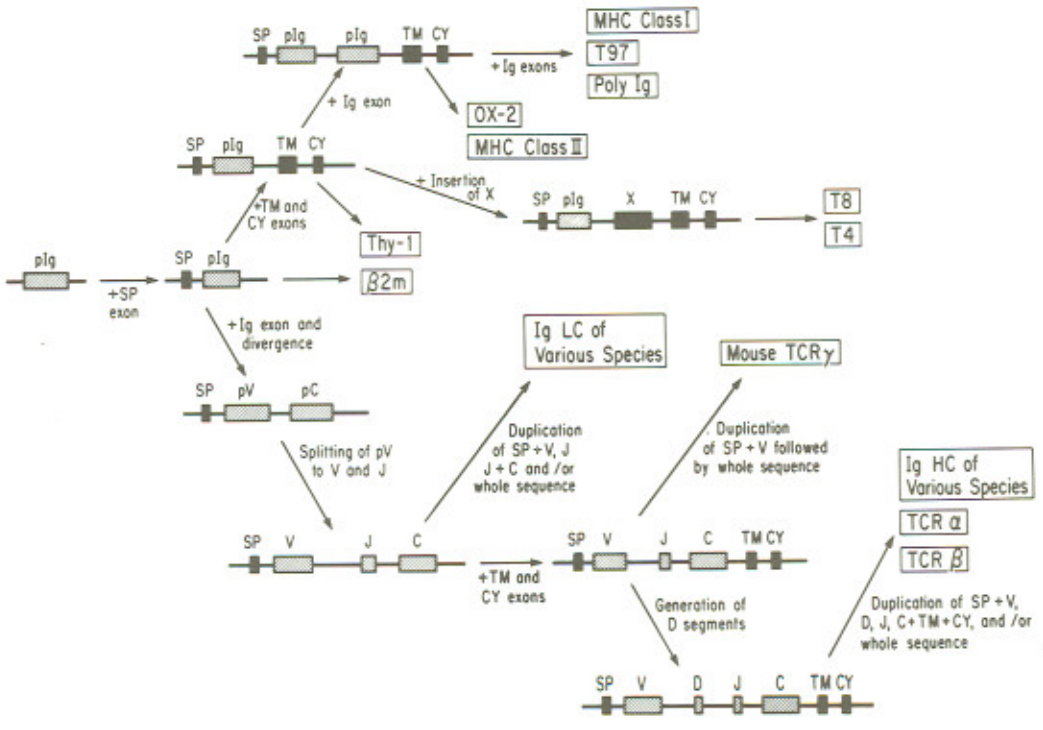


Fig. 9. A possible pathway for the evolution of immunoglobulin supergene family.

organization of the various immunoglobulin and T cell receptor gene segments and their nucleotide sequences shed light on the evolutionary origin of the somatic rearrangement: It seems that under the pressure to diversify the V region and to keep the monospecific nature of B or T cell clones, the ancestor of the vertebrates took advantage of the accidental insertion of a transposition-like element into an ancestral V opy. The present somatic *V–J* joining can be viewed as the reversal of this insertion event.

The detailed analysis of immunoglobulin and T cell receptor genes and discoveries of the ever increasing number of proteins containing immunoglobulin homology units and even immunoglobulin-like domains are contributing to our understanding of the behaviour of eukaryotic genes in general in evolution. First, the exon–intron structure of these genes clearly demonstrate the correlation of an exon and a functional and/or structural unit of the protein. Second, all immunoglobulin-like genes can be generated from a common primordial gene of one homology unit by duplication, subsequent sequence divergency of the duplicated exon, and an occasional addition of a new exon or a various combination of these events (Fig. 9). This dramatically illustrates the tremendous versatility of the gene creation mechanism based on exon shuffling. Third, at least some introns in these genes and by inference in some other genes are almost certainly derived from the spacer between the duplicated exons. Fourth, it is intriguing that most of the immunoglobulin-like proteins seem to be integral membrane proteins. This suggests that the immunoglobulin fold is repeatedly used in evolution for binding of a ligand. Alternatively, the domain–domain interaction akin to the V_L-V_H or $C_L-C_{H_2}$ may be widely used by cell surface receptors in order to accomplish interactions between cells.

Finally, the discovery of the structurally related and functionally distinct sets of proteins (sometimes referred to the immunoglobulin supergene family) make one wonder whether other eucaryotic genes may belong to their own supergene families. If so, the eucaryotic genome may be composed of no more than a few thousand supergene families.

Acknowledgements

We wish to thank many colleagues both at Basel Institute for Immunology and M.I.T. for their contribution. We also thank P. J. Maddon and R. Axel for communicating unpublished results. The work carried out at Basel Institute was supported by Hoffmann La-Roche Co. and that at M.I.T. by NIH grants CA28900 and AI17879.

References

1. Tonegawa, S., *Nature* (*London*) **302**, 575 (1983).
2. Sakano, H., Rogers, J. H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S., *Nature* (*London*) **277**, 627 (1979).
3. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. and Gilbert, W., *Proc. Natl. Acad. Sci. USA* **75**, 1485 (1978).
4. Jung, A., Sippel. A. E., Grez, M. and Schutz, G., *Proc. Natl. Acad. Sci. USA* **77**, 5759 (1980).
5. Yamada, Y., Avvedimento, V. E., Mudry, J. M., Ohkubo, H., Vogeli, G., Irani, M., Pastan, I and de Crombrugghe, B., *Cell* **22**, 887 (1980).
6. Odermatt, E., Tamkun, J. W. and Hynes, R. O., *Proc. Natl. Acad. Sci. USA* **82**, 6571 (1985).
7. Sakano, H., Hüppi, K., Heinrich, G. and Tonegawa, S., *Nature* **280**, 288 (1979).
8. Kindreed, B. and Shreffler, D. C., *J. Immun.* **109**, 940 (1972).
9. Katz, D. H., Hamaoka, T., Dorf, M. E., Maurer, P. H. and Benacerraf, B. J., *J. Exp. Med.* **138**, 734 (1973).
10. Zinkernagel, R. M. and Doherty, P. C., *Nature* (*London*) **248**, 701 (1974).
11. Zinkernagel, R. M., Callahan, G. N., Althage, A., Cooper, S., Klein, P. A. and Klein, J., *J. Exp. Med.* **147**, 882 (1978).
12. Bevan, M. J. and Fink, P. J., *Immun. Rev.* **42**, 3 (1978).
13. von Boehmer, H., Haas, W. and Jerne, N. K., *Proc. Natl. Acad. Sci. USA* **75**, 2439 (1978).
14. Allison, J. P., McIntyre, B. W. and Bloch, D., *J. Immunol.* **129**, 2293 (1982).
15. Haskins, K., Kubo, R., White, J., Pigeon, M., Kappler, J. and Marrack, P., *J. Exp. Med.* **157**, 1149 (1983).
16. Meuer, S. C., Fitzgerald, K. A., Hussey, R. E., Hogdon, J. C., Schlossman, S. F. and Reinherz, E. L., *J. Exp. Med.* **157**, 705 (1983).
17. Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S. P., Aleksander, I. and Mak, T. W., *Nature* (*London*) **308**, 145 (1984).
18. Hedrick, S. M., Nielsen, E. A., Kavaler, J., Cohen, D. I. and Davis, M. M., *Nature* (*London*) **308**, 153 (1984).
19. Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. and Tonegawa, S., *Nature* (*London*) **312**, 36 (1984).
20. Chien, Y.-H., Becker, D. M., Lindsten, T., Okamura, M., Cohen, D. J. and Davis, M. M. *Nature* (*London*) **312**, 31 (1984).
21. Novotný, J., Tonegawa, S., Saito, H., Kranz, D. and Eisen, H., *Proc. Natl. Acad. Sci. USA* **83**, 742 (1986)
22. Gascoigne, N., Chien, Y.-H., Becker, D., Kavaler, J. and Davis, M. M., *Nature* (*London*) **310**, 387 (1984).
23. Mallissen, M., Minard, K., Mjolsness, S., Kronenberg, M., Goveriman, J., Hunkapiller, T., Prystowsky, M. B., Yoshikai, Y., Fitch, F., Mak, T. W. and Hood, L., *Cell* **37**, 1101 (1984).
24. Hayday, A. C., Diamond, D. J., Tanigawa, G., Heilig, J. S., Folsom, V., Saito, H. and Tonegawa, S., *Nature* (*London*) **316**, 828 (1985).
25. Winoto, A., Mjolsness, S. and Hood, L., *Nature* **316**, 832 (1985).
26. Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N. and Tonegawo, S., *Nature* (*London*) **309**, 757 (1984).
27. Hayday, A. C. Saito, H., Gillies, S. D., Kranz, D. M., Tanigawa, G., Eisen, H. N. and Tonegawa, S., *Cell* **40**, 259 (1985).
28. Kranz, D. M. *et al.*, *Science* **227**, 941 (1985).
29. Kranz, D. M., Saito, H., Heller, M., Takagaki, Y., Haas, W., Eisen, H. N. and Tonegawa, S., *Nature* (*London*) **313**, 752 (1985).
30. Heilig, J. S., Glimcher, L. H., Kranz, D. M., Clayton, L. K., Greenstein, J. L., Saito, H., Maxam, A. M., Burakoff, S. J., Eisen, H. N. and Tonegawa, S., *Nature* (*London*) **317**, 68 (1985).
31. Raulet, D. H., Garman, R. D., Saito, H. and Tonegawa, S., *Nature* **314**, 103 (1985).
32. Collins, M. K. L., Tanigawa, G., Kissonerghis, A.-M., Ritter, M., Tonegawa, S. and Owen, M. J., *Proc. Natl. Acad. Sci. USA* **82**, 4503 (1985).
33. Strominger, J. *et al.*, *Scand. J. Immunol.* **11**, 573–593 (1980).
34. Larhammar, D. *et al.*, *Cell* **30**, 153 (1982).
35. Williams, A. and Gagnon, J., *Science* **216**, 696 (1982).
36. Mostov, K. E. *et al.*, *Nature* (*London*) **308**, 37 (1984).
37. Kavathas, P., Sukhatme, V. P., Herzenberg, L. A. and Parnes, J. R., *Proc. Natl. Acad. Sci. USA* **81**, 7688 (1984).
38. Littman, D. R., Thomas, Y., Maddon, P. J., Chess, L. and Axel, R., *Cell* **40**, 237 (1985).
39. Maddon, P. J., Littman, D. R., Godfrey, M., Maddon, D. E., Chess, L. and Axel, R., *Cell* **42**, 93 (1985).
40. Clark, M. J., Gagnon, J., Williams, A. F. and Barclay, A. N., *EMBO J.*, **4**, 113 (1985).
41. Tonegawa, S., *Scientific American* **253**, 122 (1986).
42. Novotný, J. and Auffray, C., *Nucleic Acids Res.* **12**, 243 (1984).
43. Segal D., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M. and Davies, D., *Proc. Natl. Acad. Sci. USA* **71**, 4298 (1974).