beled μ probe, or in some cases, with δmRNA that was labeled with ^{32}P by using polynucleotide kinase as described by N. Maizels, *Cell* **9**, 431 (1976).
26. See the article by P. W. Tucker, C.-P. Liu, J. F. Mushinski, F. R. Blattner, *Science* **209**, 1353 (1980).
27. N. M. Gough *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 554 (1980).
28. To calculate this, we first calculated the distance from the Xba I site in the 3′ untranslated region of the μ membrane exon to the Bgl II site just to the left of Cδ1 by subtracting the coordinates (*22*) (6618 − 4457 = 2161). We then added to this value 145 bp for the Bgl II to Cδ1 distance and 160 bp for the distance from the Cμ exon terminator to the Xba I site (*7*). The result was 2466 nucleotide pairs.
29. S. Slavin and S. Strober, *Nature (London)* **272**, 624 (1978); M. R. Knapp, E. Severison-Gronowicz, J. Schroeder, S. Strober, *J. Immunol.* **123**, 1000 (1979); M. Knapp, personal communication regarding the observation that IgM and IgD molecules on BCL-1 lymphoma cells show the same idiotype.
30. W. J. Dreyer and J. C. Bennett, *Proc. Natl. Acad. Sci. U.S.A.* **54**, 865 (1965).
31. We thank M. Fiant and J. Richards for work on heteroduplex and electromicroscopy, E. Lui for technical help on phage plaque purification, and W. Fiske and E. La Luzerne for manuscript preparation. All experiments were carried out in accordance with the NIH Guidelines on Recombinant DNA Research. Supported by grants GM 21812 (F.R.B.) and GM 06768 (C.P.L.) from the National Institute of General Medical Sciences, and by grant AI 16547 (P.W.T.) from the National Institute of Allergy and Infectious Diseases. This is paper number 2447 from the Laboratory of Genetics, University of Wisconsin.

16 July 1980

Mouse Immunoglobulin D: Messenger RNA and Genomic DNA Sequences

Philip W. Tucker, Chih-Ping Liu
J. Frederic Mushinski, Frederick R. Blattner

Summary. The molecular structure of a mouse immunoglobulin D from a plasmacytoma tumor and that of the normal mouse gene coding for immunoglobulin D are presented. The DNA sequence results indicate an unusual structure for the tumor δ chain in two respects: (i) Only two constant (C) region domains, termed Cδ1 and Cδ3 by homology considerations, are found; the two domains are separated by an unusual hinge region CδH that lacks cysteine residues and thus cannot provide the covalent cross-links between heavy chains typically seen in immunoglobulins. The two domains and hinge are all coded on separate exons. (ii) At the carboxyl end of the δ chain there is a stretch of 26 amino acids that is coded from an exon located 2750 to 4600 base pairs downstream from the rest of the gene. Analogy with immunoglobulin M suggests that this distally coded segment CδDC may have a membrane-binding function; however, it is only moderately hydrophobic. A fifth potential exon (CδAC), located adjacent to the 3′ (carboxyl) end of Cδ3, could code for a stretch of 49 amino acids. The tumor's expression of the δ gene may be aberrant, but the simplest interpretation would be that this tumor expresses one of the several biologically significant forms of the δ chain.

Immunoglobulin molecules on the membranes of lymphocytes serve as receptors and in this form play a pivotal role in regulating the immune response of an animal to antigenic challenge (*1*). One example is the antigen-dependent triggering of B cells (*2*). To be eligible to differentiate into a proliferating antibody-secreting cell, a particular B cell must produce and bear on its surface an immunoglobulin that can react with the antigen. Thus the animal's protein-synthesizing resources are not devoted to large-scale synthesis of useless antibodies for which an antigen is not present.

Superimposed on this regulatory process are additional controls that prevent triggering of B cells bearing self-reactive antibodies (*3*) and, in the case of cellular antigens, histocompatibility restrictions that regulate T cell–B cell interactions (*4*, *5*).

Immunoglobulin D (IgD) whose role is at present poorly understood, is a major immunoglobulin receptor on B cell membranes (*6*). The IgD class is distinguished by its δ type heavy chains. Membrane expression of IgD may be associated with such functions as memory propagation (*7*) and maturation (*8*), prevention of tolerance (*9*), idiotype suppression (*10*), and early B cell development (*11*). Although IgM appears first in ontogeny (*12*, *13*), IgD is present along with IgM on a majority of splenic B lymphocytes (*12*, *14*). On each cell these two classes of receptor share the same antigenic specificity and idiotype (*15*), which probably means that they have the same variable region.

Most of the available knowledge about immunoglobulin structure has come from analysis of plasmacytoma proteins, which are assumed to reflect the secreted rather than the membrane form of antibodies. Both types of immunoglobulin molecules are generally composed of paired heavy (H) and light (L) chains, each of which is made up of a series of domains of about 110 amino acid residues. Both heavy and light chains contain a variable (V) region domain at the amino terminal end; the heavy chain and light chain V regions interact to form the antigen-combining site. Light chains have a single constant (C_L) region domain located at the carboxyl terminus which is linked to the first constant region domain of the heavy chain (C_H1). In all normal secreted immunoglobulins so far studied, heavy chains have either three or four C_H domains depending on class. The heavy chains interact with each other covalently by disulfide bonds and noncovalently by other interactions to form an H_2L_2 tetramer (which in immunology is traditionally termed a monomer) of characteristic three-dimensional structure (*16*). In IgM and IgA, the secreted form of these monomers are further polymerized. In the IgG class, a peptide segment termed the hinge is present between the first and second C_H domains. This segment of amino acids is generally rich in proline, is involved in disulfide bonding between heavy chains,

Philip W. Tucker was assistant professor, Department of Biochemistry, University of Mississippi Medical Center, Jackson 39216. He is now associate professor, Department of Microbiology, University of Texas Southwestern Medical School, Dallas 75235. Chih-Ping Liu is a postdoctoral fellow, Laboratory of Genetics, University of Wisconsin–Madison, Madison 53706. J. Frederic Mushinski is senior investigator, Laboratory of Cell Biology, National Cancer Institute, Bethesda, Maryland 20205. Frederick R. Blattner is associate professor of genetics, Laboratory of Genetics, University of Wisconsin–Madison, Madison 53706.

0036-8075/80/0919-1353$01.75/0

and permits a high degree of segmental flexibility for the antigen-binding arms, thus allowing an optimal fit of the combining site to antigenic determinants at varying distances and angles on the antigens (*16*).

Recent studies done by DNA cloning techniques have shown that immunoglobulin genes are expressed from a number of separate "exonic" segments of DNA that are arranged in modules in the genome (*17*). Several types of rearrangements, involving both DNA joining and RNA splicing, occur under developmental control during the expression of immunoglobulins (*18*), and it is becoming increasingly attractive to speculate that "exon shuffling" plays a major role in the regulation of the immune response.

Recently, it has become possible for some deductions about immunoglobulin proteins to be made more easily and rapidly from nucleic acid studies than from studies of the proteins themselves. An example of this is the work of Early *et al.* (*19*) and Rogers *et al.* (*20*) in defining a carboxyl terminal exon shuffle (see discussion by Liu *et al.*, pages 1348–1353) that may be responsible for the secreted and membrane forms of IgM; their findings were consistent with the analysis of the secreted and membrane forms of the μ protein chains (*21*).

A carboxyl terminal exon shuffle may not be the complete explanation in all cases for differences between membrane-associated and secreted forms of immunoglobulins. For example, Oi *et al.* (*22*) identified antigenic differences between secreted and membrane forms of IgG2a localized to the hinge or $C\gamma_{2a}2$ domain. Such changes could also be mediated by exon shuffles.

For IgD, data on the protein have been particularly hard to come by. Analysis of the secreted form of the IgD protein has been difficult because of its low abundance in serum, its extreme susceptibility to proteolytic cleavage, and the rarity of IgD myelomas (*23*). Nonetheless, Spiegelberg (*24*) succeeded in obtaining limited amino acid sequence from a human IgD myeloma Fc fragment and concluded that IgD showed some resemblance to IgG. These findings have recently been corroborated and extended by Lin and Putnam (*25*) on another human IgD myeloma protein, WAH. The human δ chain appears to have three typical C_H domains and an extended hinge of some 50 residues between Cδ1 and Cδ2.

Mouse serum has an extremely low concentration of IgD. Two sizes of δ chain from serum, δ1 and δ2, have been estimated as 57,000 and 69,000 daltons, respectively (*26, 27*). The membrane form of IgD has been characterized in several studies (*28–30*). These reports, which rely on labeling of membrane preparations and specific immunoprecipitation techniques, generally agree that the membrane δ chain is a glycoprotein of 65,000 to 72,000 daltons, as determined by gel electrophoresis. However, all the measurements of δ chain size radiolabeled in vivo are strongly affected by glycosylation. By one estimate (*29*) the length of the δ chain after enzymatic carbohydrate removal was 50,000 daltons, which is about 20,000 daltons less than the fully glycosylated protein. The major but uncertain contribution of carbohydrate to molecular size has hampered attempts to determine the size of the δ chain and the organization of its domains. An additional finding that may have biological significance is that there may be IgD "half-molecules," consisting of only a single light and a single heavy chain on lymphocyte surfaces (*31, 32*), in addition to the standard H_2L_2 monomer.

One might hope to obtain an understanding of IgD function by careful analysis of its protein structure and its genomic organization. The article by Liu *et al.* (*33*) showed the relationship of the Cμ and Cδ genes in the mouse genome. In this article, we focus on the fine structure of Cδ exonic segments.

Materials and Methods

The recently classified tumors, TEPC 1017 and TEPC 1033 (*34*), are the first mouse plasmacytomas that have been characterized as producing an immunoglobulin of the IgD class. Immunoglobulin D, which is identified by immunoprecipitation by the 10-4.22 hybridoma antibody (*35*), is both secreted from and found on the surface of these cells (*34*).

Polyadenylated RNA was isolated from TEPC 1017 and was used (*36*) to direct the cell-free synthesis of a 44,000-dalton protein specifically immunoprecipitable by a class-specific antibody to δ (*34*). From this heavy chain messenger RNA (mRNA), a complementary DNA (cDNA) plasmid clone, pδ54J, was constructed by a modification of the method of Maniatis *et al.* (*37*). The construction of this plasmid and its verification as an authentic IgD clone by mRNA selection will be described elsewhere (*36*).

The pδ54J cDNA plasmid was used in (*33*) to isolate genomic clone Ch28 257.3, which contains the δ gene from BALB/c mouse liver DNA. This germ-line clone and the cDNA clone were used for the structural studies described in this article.

Arrangement of Cδ Exons in BALB/c DNA

The first step of analysis of the Cδ gene was the preparation of fine-scale maps of the cloned δ chain mRNA in pδ54J and the corresponding BALB/c genomic DNA cloned in Ch28 257.3. We used agarose gel electrophoresis, Southern hybridization (*38*), and DNA sequencing techniques for this purpose. Figure 1 shows the results of the mapping. Dotted lines are used to indicate the relationship between genomic and mRNA segments. We found that the TEPC 1017 δ constant region mRNA is derived from four exons in genomic DNA. The first three exons code for two constant region domains separated by a hinge segment and map into the leftmost of the two hybridizing regions identified in (*33*). The fourth exon codes for the 26 amino acids at the carboxyl terminal end of the δ chain and maps to the rightmost hybridizing region. This distally coded segment was mapped in the genomic DNA clone by Southern hybridization with the 3′ terminal Hinf I–Pst I fragment of pδ54J as a probe. Since we have not refined the mapping in this area yet, we can only place this exon within the region between 2750 and 4600 base pairs downstream from the rest of the gene. We also cannot exclude the possibility that the distal segment is coded on more than one exon within that area.

Nomenclature

The arrangement of the mouse Cδ gene is typical of other immunoglobulin genes in that domains and hinge segments are coded by separate exons. However, the presence of only two domains separated by a hinge in the Cδ constant region is unprecedented. Other heavy chains, including the human δ chain WAH, have at least three C_H domains. We do not yet know whether the short δ chain of TEPC 1017 is an aberrant arrangement, but it clear from the homology consideration discussed below that the second domain we found is more like the human Cδ3 domain than the human Cδ2. We have therefore decided to call the TEPC 1017 mouse Cδ segments

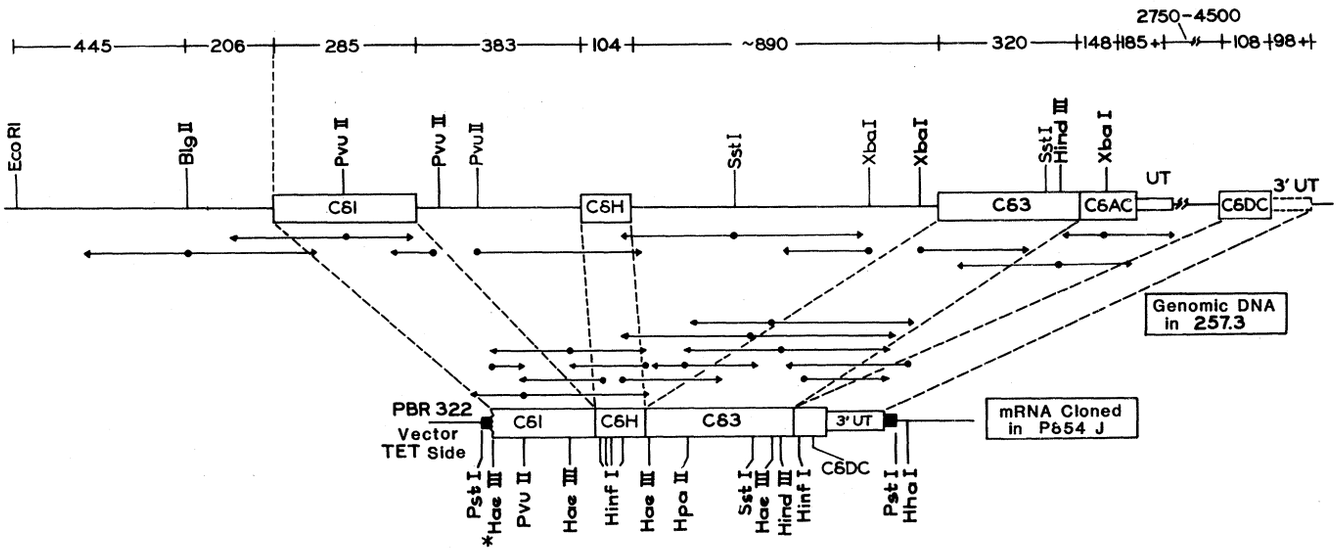


Fig. 1. Relationship of Cδ gene and TEPC 1017 δ mRNA. The region of the BALB/c genome cloned in Ch28 257.3 (*33*) that contains the Cδ gene is shown in the upper panel. Distances are given in base pairs. Exons are designated by boxes, and introns and flanking regions are designated by lines. The restriction map for the cDNA insert of pδ54J cloned from TEPC 1017 tumor mRNA is given in the lower panel. The beginning of the insert in the plasmid is nucleotide 90. Dashed lines from the genomic map indicate the contiguous exonic arrangement in the mRNA of Cδ1, CδH, Cδ3, and CδDC segments and 3′ untranslated region (*3′ UT*). For both DNA's, the dots and arrows denote the DNA sequencing strategy. Fragments were cleaved and labeled at the dots, and the direction and length of the sequence obtained from these sites, using the Maxam and Gilbert method (*39*), are indicated by the arrows.

Cδ1, CδH, and Cδ3, rather than numbering the domains consecutively as would be the usual practice for a new plasmacytoma protein.

The carboxyl terminal arrangement of the δ gene also presents a problem of nomenclature. This arrangement is very similar to that of Cδ, in which the membrane (M) terminus, μ_M, is coded distally to the last domain and the secreted (S) terminus, μ_S, is coded adjacent to the last domain. However, for Cμ, information was available to indicate a membrane-binding function for the distally coded terminus, and protein sequence data proved that the adjacently coded segment is used to code the secreted form of the antibody. Neither of these relationships has been shown for Cδ. Since it would be unwise to assume that topographical similarity defines functional analogy, we refer to these regions in the δ gene as CδAC (adjacently coded) and CδDC (distally coded) instead of as S and M.

Amino Acid Sequence of Mouse δ Chain as Deduced from DNA Sequence

The structure of the TEPC 1017 protein indicated by our analysis of both mRNA and gene is unusual for an immunoglobulin heavy chain in that only two constant region domains separated by a hinge are present. Before discussing the possible significance of these findings, we will comment on the sequences themselves.

Figure 2 shows the DNA sequence [see (*39*)] obtained for each exonic segment of the δ gene, including the CδAC segment, which may at different times be either intronic or exonic, if the expression of this gene follows the pattern of the μ gene. The data are presented with 5′ sides of the segments aligned to facilitate comparisons. All of the sequences presented in Fig. 2 were determined from both strands of both mRNA and BALB/c genomic DNA with three exceptions: (i) 95 nucleotides at the 5′ end of Cδ2 were determined from the two strands of genomic DNA only; (ii) CδAC was sequenced from the two strands of genomic DNA only; and (iii) CδDC was sequenced only from the two strands of the mRNA clone.

Translation of the DNA sequences allowed the amino acid sequence to be predicted (see Fig. 2). We found that the Cδ1 and Cδ3 mouse amino acid sequences have structures expected for immunoglobulin domains (*16*) with conserved Cys (*40*) residues (denoted by $ in Fig. 2) that enclose approximately 60 amino acids and include conserved Trp residues (in Fig. 2) within the disulfide loops. The Cδ1 domain exhibits generic properties characteristic of C_H1 domains of heavy chains. For example, all C_H1 domains sequenced to date, except for two, have one rather than two invariant Trp residues located on the carboxyl terminal side of the first invariant Cys residue (*41*). The intradomain disulfide loop in C_H1 domains is also characteristically longer (average, 66 residues including the Cys residues) than that in C_H2 domains (average, 53 residues) (*41*). Furthermore, all mouse C_H1 domains have in addition to their invariant Trp and Cys residues several other highly conserved amino acids (* in Fig. 2). All of these criteria are met by the domain that we have termed Cδ1. Potential linkage to the light chain may be provided by either of the two Cys residues on the amino terminal side of that proposed for the intra-Cδ1 domain loop.

Since the cDNA clone pδ54J is not long enough to include the 5′ terminus of Cδ1, the choice of this 5′ boundary is based on the identification of an RNA splice site in the genomic DNA sequence flanking Cδ1. We have chosen the third of four feasible 5′ (acceptor) RNA splice sites. This junction is more compatible with the consensus donor derived on statistical grounds (*42*) (see Fig. 4) and fits the U1 RNA sequence (*42*) better. The other possibility is indicated by a slash in Fig. 1. Our preferred choice provides a Cδ1 domain of 102 amino acids, which is similar to other C_H1 domains (*41*).

Amino acid sequence homology comparisons were made between the Cδ1 domain and the corresponding C_H domains for which complete sequence information is available (*41*). The method is to introduce gaps when necessary into the

Fig. 2. Nucleotide sequence of Cδ exons. The segments denoted Cδ1, CδH, and Cδ3 were sequenced both from genomic (Ch28 257.3) and TEPC 1017 tumor mRNA (pδ54J), whereas CδAC was determined exclusively from genomic clones and CδDC was determined exclusively from cDNA. Refer to the text for the significance and continuity of these segments. The symbol + appears under every tenth base. The slash at nucleotide 36 in Cδ1 indicates a potential RNA splice site (see Fig. 4). Amino acids predicted to be translated are directly below the first base of the corresponding codon. The domain sizes in amino acid residues/base residue are: Cδ1, 102/304; CδH, 35/105; Cδ3, 107/321; CδAC, 49/321; CδDC, 26/206.

sequences in order to align the invariant intradomain Cys and Trp residues and other regions of homology (note * in Fig. 2). The gaps correspond to single evolutionary deletion or insertion events. The Cδ1 domain was most similar to the $C\gamma_1 1$ domain (26 of 93 residues, ≈ 28 percent agreement, with four gaps including 11 amino acids) and somewhat more similar to the other γ subclass domain ($C\gamma_{2b}1 \approx C\gamma_{2a}1$, which is ≈ 23 percent, with four gaps, including 8 amino acids) than to either Cα1 (15 of 87 residues, ≈ 17 percent, with three gaps including 15 residues) or Cμ1 (17 of 90 residues, ≈ 19 percent, with three gaps including 11 residues). Since it is generally assumed that all C_H domains have evolved by duplication and subsequent mutation from a common ancestral sequence, the Cδ1 and $C\gamma_1 1$ domains appear to be more recent relatives. Similar conclusions were reached for the human by Spiegelberg (*24*) after analysis of the first 53 residues of a myeloma Fc fragment.

The Cδ hinge region, as with mouse γ_1 and γ_{2b} chains (*43*), is encoded on a separate exon (CδH in Fig. 1). This 35-amino acid segment is longer than any other sequenced mouse hinge, yet in contrast with other hinge sequences, no Cys residues are found. Thus no disulfide linkages between heavy chains are possible for this region of the protein. The CδH region contains only one Pro residue, which is also unusual since most hinges are proline-rich (*41*). Lin and Putnam (*25*) have reported fragmentation and composition results indicating that the IgD protein WAH hinge is about 50 residues long, rich in *O*-linked sugars and unexpectedly low in Pro residues. The WAH hinge is highly charged at its carboxyl terminus because of a cluster of seven Lys, three Arg, and ten Glu residues (*44*). The mouse δ hinge is also relatively highly charged with six basic and five acidic residues out of 35 (Fig. 2), but these do not appear to be clustered. There are also five Ser and six Thr residues which could serve as sites for *O*-glycosyl linkage of carbohydrate moieties.

The Cδ3 domain has generic properties unique to carboxyl terminal domains of immunoglobulins. The disulfide loop ($ in Fig. 2) contains 60 amino acids with two invariant Trp residues included (& in Fig. 2). Residues in the region of the invariant Cys at the carboxyl terminal end are similar to all carboxyl terminal domains (* in Fig. 2), and the homology of Cδ3 with Cμ4 is particularly strong in that area (9 of 10 residues). The amino acid sequence of mouse Cδ3 matches the Cδ3 domain of WAH in 52 percent of the positions up to the end of the Cδ3 exon in mouse genomic DNA. Beyond this point the WAH protein extends for seven amino acids, only one of which matches Cδ3 (*44*).

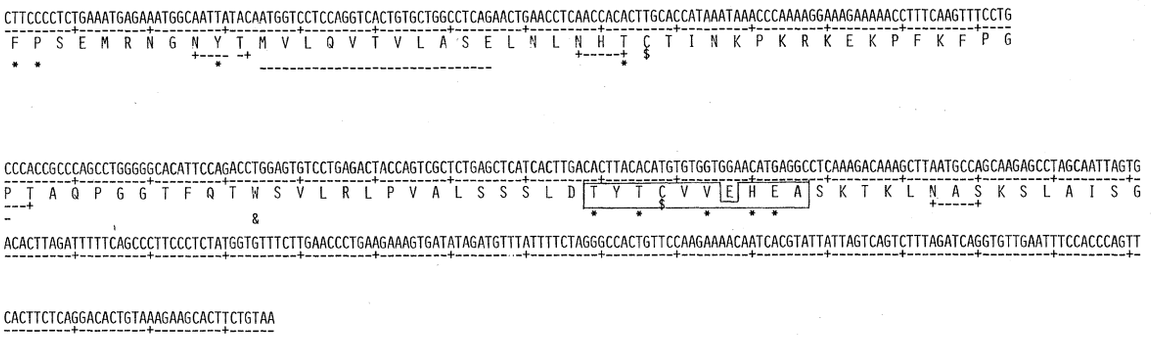
Potential Alternate Carboxyl Termini for the δ Chain

The CδDC exonic segment codes for the 26 amino acids at the carboxyl terminal end of the δ chain coded by TEPC 1017 mRNA, whereas the CδAC segment would code for an alternate carboxyl terminus of 46 amino acids if it were transcribed into RNA. Figure 3 shows a comparison of these alternative carboxyl termini with the two analogous ones for IgM. We computed several hydrophobicity indices and secondary structure predictions for the 26-residue CδDC peptide and for a 12-residue subfragment of this sequence denoted CδDC′. These values were compared to the 26-amino acid hydrophobic core of μ_M, to the entire 49 residues of CδAC peptide, and to the hydrophobic core of the prototypic transmembranal protein, glycophorin (*45*). As anticipated, the CδAC terminus is very polar, and it is unlikely that this segment could interact with a membrane unless the buried charges were neutralized by ion pair formation. Models for such interactions have been proposed (*46*).

The calculations for CδDC do not clearly suggest that this is a transmembrane protein, nor do they rule it out. The peptide is moderately hydrophobic, even though the presence of three acidic residues elevates the polarity values considerably. The stretch of amino acids between charges (CδDC′) is extremely nonpolar and it may be long enough to span a lipid bilayer in a β-pleated sheet conformation (*47*). On the other hand, the core structure of μ_M is extremely hydrophobic as judged by all criteria. Rogers *et al.* (*20*) proposed that this segment was a transmembranal α-helix with its positively charged carboxyl terminus providing the link to the cytoplasm. We calculate that the μ_M core could span the membrane equally well either on an α-helical or a β-pleated sheet conformation. Very similar parameters are observed for glycophorin. We conclude that a membrane association for the CδDC structure is not beyond reason according to present calculations, but the parameters place this segment within the range of normal nonmembrane associated segments of amino acids (*48*). It appears that if CδDC does function in binding the lymphocyte membrane, it probably does so in a manner different from that of μ_M.

Carbohydrate Attachment Sites

The sequenced portion of the δ protein of TEPC 1017 has seven instances of Asn-X-Thr or Asn-X-Ser (where X can be any amino acid) (see Fig. 2), which can serve as a recognition site for carbohydrate addition via the dolichol phos-



Invariant Cys residues proposed to participate in intradomain linkages are indicated with $. Conserved Trp residues are marked by &, and other conserved amino acids that are diagnostic of domain type (note discussion in text) are indicated by *. The residues boxed in Cδ3 match the corresponding region of the MOPC 104E μ chain (*44*). Proposed *N*-linked sugar attachment sites are indicated by +- - -+. The single letter abbreviations are *A*, Ala; *C*, Cys; *D*, Asp; *E*, Glu; *F*, Phe; *G*, Gly; *H*, His; *I*, Ile; *K*, Lys; *L*, Leu; *M*, Met; *N*, Asn; *P*, Pro; *Q*, Gln; *R*, Arg; *S*, Ser; *T*, Thr; *V*, Val; *W*, Trp; *Y*, Tyr [see (*40*)].

phate *N*-linked pathway (*49*). It is not surprising that measurements of the size of IgD chains in gels have been heavily influenced by carbohydrate moieties that decrease electrophoretic mobility.

RNA Splice Sites

The DNA sequences that span the intron-exon junctions of the Cδ gene are presented in Fig. 4. As in other eukaryotic genes, potential RNA splice sites can be chosen in each case to obey the GT-AG rule (*50*), and the neighboring intronic sequences share features observed in the consensus splice sequences (*42*) denoted in Fig. 4. Of the four potential junctions for the 5′ side of the Cδ1 domain, the second appears by homology considerations to be the best choice. A consensus 5′ (donor) splice site appears at the end of the Cδ3 domain between the dipeptide Gly-Lys in the putative CδAC form of the protein. This position is the same in Cμ and, as pointed out by Rogers *et al.* (*20*), corresponds to the ends of both IgG class heavy chains that have been sequenced at the DNA level (*43*). This strengthens the analogy between Cμ and Cδ in this area.

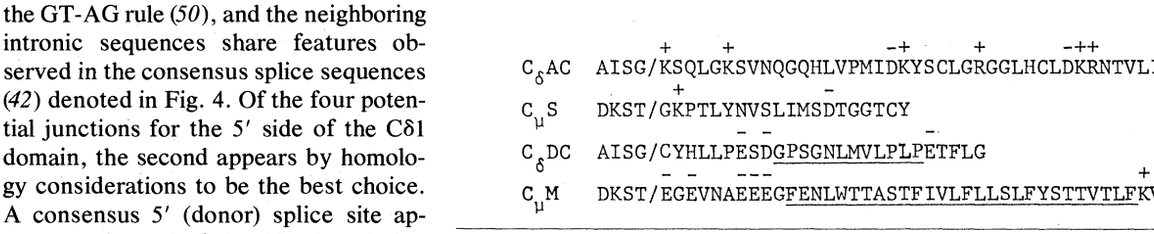
3′ Untranslated Regions

The 3′ untranslated region of TEPC 1017 δ mRNA as cloned in pδ54J is shown in Fig. 1. Its sequence (Fig. 2) extends from the terminator at the end of CδDC for 125 nucleotides before reaching the homopolymeric polycytidylate tails (not shown) used for the plasmid construction. The absence of the prototype polyadenylate addition sequence AATAAA (*51*) suggests that we have not cloned all of the 3′ untranslated region. We have also sequenced approximately 170 base pairs rightward from the termination codon of the CδAC exon (Fig. 2) without reaching any putative polyadenylate addition site.

Is What We Sequenced the Major δ Chain Expressed by TEPC 1017?

Several lines of evidence converge to show that the sequence we have determined corresponds to the major heavy chain expressed in TEPC 1017. Rudikoff (*52*) has sequenced two 11-amino acid stretches from the amino termini of cyanogen bromide fragments of the TEPC 1017 δ chain. These sequences matched segments of both Cδ domains underlined in Fig. 2. This showed that the mRNA and gene segments we have sequenced code for amino acid sequences present in the major IgD protein produced by the plasmacytoma.

From the amino acid composition of Cδ1, CδH, Cδ3, and CδDC we calculate a molecular size of 30,040 daltons for the constant region of the TEPC 1017 δ



Parameter	Carboxyl terminal segments				
	CδAC	Cμ M core	CδDC	CδDC′	Glycophorin
Polarity	14.2	0.947	8.5	0.985	0.432
Internal residue preference	0.977	1.268	1.123	1.223	1.637
Hydrophobicity (ΔG) (surface tension)	−820	−1466	−1110	−1100	−1447
Hydrophobicity (ΔG) (transfer)	−612	−1323	−719	−692	−1495
α-Helix potential (P_α)	0.959	1.039	0.971	.912	0.973
β-Sheet potential (P_β)	1.045	1.185	0.952	1.034	1.295

Fig. 3. Comparison of alternative carboxyl terminal amino acid sequences predicted from δ and μ chain DNA sequences. The upper comparison is a sequence alignment in which the junction of the last domain (Cδ3 or Cμ4) is indicated by the slash. Basic (+) and acidic (−) residues are indicated, and the underlined regions of the CδDC and CμM are referenced below as the CδDC′ and CμM cores, respectively. The comparison below gives hydrophobicity and secondary structure predictions for these sequences and for the membrane-binding segment of glycophorin (*45*). All parameters have been normalized to a per residue basis and are defined as follows. Polarity (*46*) reflects the number of charged residues and tendency of uncharged residues to be polarized; larger values indicate greater polarity. The internal residue preference (*62*) predicts the ability of peptide to be buried in a nonpolar environment; larger values indicate greater preference. The hydrophobicity-surface tension method (*63*) is based on the transfer of amino acid from the aqueous solution to the surface phase; ΔG is the change in free energy in kilocalories per mole; more negative values indicate greater hydrophobicity. Hydrophobicity-transfer (*64*) is based on the transfer of residues from water to alcohol. The Chou-Fasman secondary structure predictions (*65*) indicate that residues with the highest P_α values will initiate α-helix formation and those with the highest P_β values will initiate β-pleated sheet formations; values > 1.000 indicate a significant tendency to form the structure.

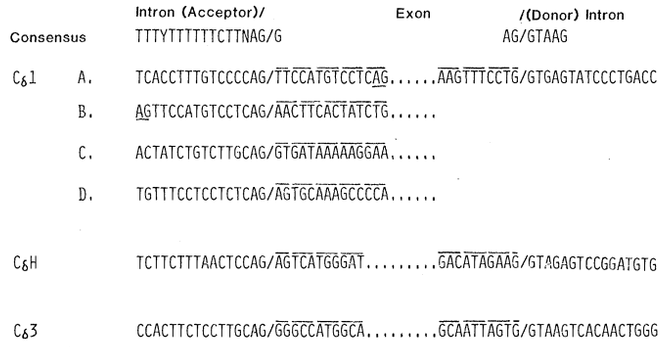


Fig. 4. Intron-exon sequence junctions in the mouse Cδ gene. Sequences spanning the RNA splice sites are aligned. The slashes denote splices that follow the GT-AG rule (*50*) without altering protein sequence. Assignment of the 5′ boundary of Cδ1 is tentative, and four possible splice sites are denoted *A* through *D*; *B* is the preferred site used in Fig. 2, and the slash corresponds to the 5′ boundary of Cδ1 shown in Fig. 2. The intron of *B* and all of *A* are not shown in Fig. 2, since they are to the left of what we feel is the Cδ1 exon. Coding triplets are shown by horizontal lines. The consensus RNA splice sites are from Rogers and Wall (*42*); *N* represents any base and *Y* represents either pyrimidine. The term donor refers to the site on the 5′ side of an intron and acceptor refers to the 3′ side. The overlap of *A* and *B* is underlined.

chain. Addition of 14,000 daltons for a typical V_H region and 1,000 daltons for an amino terminal leader segment of 10 to 15 amino acids (*53*) can account well for the 44,000-dalton δ chain synthesized in vitro and the 43,000-dalton δ chain made in vivo in the presence of tunicamycin (*36*), an inhibitor of *N*-linked glycosylation. Thus we feel reasonably confident that we can account for the complete δ chain of TEPC 1017 by the DNA sequences presented in this article.

Are Minor δ Chains Expressed in TEPC 1017?

Several experiments have been done to test whether additional Cδ exons are expressed in TEPC 1017. Southern hybridization experiments using labeled tumor mRNA (uncloned) as probe revealed no hybridization to regions of cloned genomic DNA other than the four exonic segments we have sequenced. Specifically, no hybridization was found to fragments on the left of the Bgl II site in Fig. 1 [see also figures 2 and 3 in (*33*)]. Negative results were found for the large Sst I–Xba I fragment derived from the intron between the CδH and Cδ1 domains and for the Xba I fragment in the genomic clone that contains CδAC. We conclude that no large amount of mRNA is present that could code for a domain to the left of Cδ1, for a Cδ2 domain from the intron, or for an adjacently coded carboxyl terminal segment.

This general conclusion was confirmed by electron microscopy of about a dozen R-loop molecules (data not shown), formed between the isolated TEPC 1017 δ chain mRNA and the genomic Cδ clone Ch28 257.4. The results are consistent with the expression of only two full-sized domains and clearly showed the large 890-base pair intron.

There are, however, some indications for alternative forms. Mushinski *et al.* (*36*), using RNA blotting procedures, noted several minor species of Cδ-hybridizing RNA from TEPC 1017. The significance of these forms remains to be determined, since it is not known which sequences are present in these molecules and whether or not they are expressed as protein. Another indication of possible minor forms is the finding that there is a significant difference between the electrophoretic mobility of membrane-labeled δ chains and secreted δ chains for TEPC 1017 (*34*). However, this result could be due to differences in glycosylation just as well as to changes at the peptide level. As a general conclusion we find that the form of TEPC 1017 δ chain that we have sequenced is the major form synthesized by this tumor, but expression of other forms in small quantities cannot be ruled out at the present time, especially if they have a structure similar to the major form.

Is Native Membrane Cδ Like TEPC 1017 Cδ?

The TEPC 1017 plasmacytoma protein should have a role in revealing the structure of δ chains that have biological significance for the immune system. However, it is difficult at present to know to what extent our results, arising as they do from a tumor, can be extrapolated to normal δ chains in their several forms on B cell membranes (*32*) and in serum (*26, 27*).

A most important initial question is whether native forms of mouse IgD have only two Cδ domains. Estimates of molecular size for membrane δ range from 62,000 to 72,000 daltons, but these estimates include a large component of unknown magnitude due to carbohydrate attachment.

The seven *N*-linked glycosylation sites that we find in the Cδ sequences could account for the large shift in molecular size (~ 20,000 daltons) accompanying removal of these moieties (*29*) if each carbohydrate adds about 3000 daltons to the apparent molecular size determined by sodium dodecyl sulfate (SDS) gel analysis.

Kessler and Finkelman (*54*) have compared radioactively labeled membrane δ chain from TEPC 1017 directly with comparable material from splenic B cells in parallel channels of a single SDS gel. They found no difference in mobility (68,000 apparent molecular weight), indicating that the membrane δ of TEPC 1017 and that of native B cells are the same size. A very large compensatory difference in carbohydrate would be needed for the two chains to differ by a domain at the peptide level, yet show the same molecular size.

A molecular weight estimate has been obtained by Goding and Herzenberg (*29*) for the native form of mouse membrane δ chain after enzymatic removal of *N*-linked core oligosaccharides. The size they found was about 50,000 daltons, and it led them to postulate three constant region domains for Cδ instead of the four that had been previously suggested. However, this measurement is only 6000 daltons (half of a domain) greater than our estimate for the TEPC 1017 protein based on sequence. The gel measurement could be an overestimate, especially since *O*-linked oligosaccharides present in the hinge, and elsewhere, may not have been removed by the enzymatic treatment used. Thus, taken as a whole, the available evidence on size is consistent with the two domains plus hinge structure for the constant region of δ chains on the membranes of mouse B cells.

A significant question concerning the relevance of the TEPC 1017 protein to membrane-bound IgD is raised, however, by the unconvincing membrane-binding potential of the CδDC segment compared with the membrane-binding component of the μ chain. This correlates with the finding that TEPC 1017 secretes IgD as well as exhibiting it on the membrane, but it is not very satisfying as an explanation for the membrane binding of native IgD, which is quite stable (*34*). Another component such as an Fc receptor may be involved in attachment of IgD to B cells. However, there is no direct

evidence that the carboxyl terminus we have sequenced is on the membrane. Instead it may serve a role that we do not yet appreciate or no role at all.

A Proposal to Explain IgD Half-Molecules

One of the striking aspects of the sequence in Fig. 2 is the lack of Cys residues that could cross-link the heavy chains to form an immunoglobulin monomer with two heavy and two light chains. The only plausible candidate for such a cross-link in our sequence would be the first residue of the CδDC segment. The other two Cys residues are in Cδ1 and are probably involved in attachment of light chains. They are certainly not involved in cross-linkage between heavy chains since it has been shown by tryptic fragmentation that all H-H cross-links are on the carboxyl terminus side of the hinge in membrane δ chains (*30*).

If the only cross-linkable Cys residue is on the CδDC segment, it might provide a mechanism to explain the finding of Eidels (*31*) and of Mescher *et al.* (*32*) of two noninterconvertible forms of surface IgD, one of which is cross-linked to form H_2L_2 monomers and the other which exists as H-L "half-molecules." To explain these two forms, alternative carboxyl terminal exons, one with a Cys residue (CδDC) and the other without, could be postulated. The same effect could also be accomplished by more elaborate Cδ domain switches.

Is There a Cδ2 Exon in BALB/c Mouse?

A major puzzle is raised by the absence of a Cδ2 domain in the TEPC 1017 protein. This is clearly in contrast to the secreted form of human IgD exemplified by WAH. We suspect that the gene for Cδ2 may be missing from the BALB/c mouse genome. The incomplete nucleotide sequence we have derived for the introns (arrows in Fig. 1) have been examined, and no obvious domain-like features or homology with the human Cδ2 amino acid sequence of the human δ chain of WAH were revealed (*44*). If in fact the mouse turns out not to have a gene for a Cδ2, it would signify a major difference between the immune systems of mouse and man. On the other hand, if a Cδ2 domain is found in the intron between CδH and Cδ3, there will have to be some explanation for its lack of expression in TEPC 1017. This lack could simply be due to a defect of the tumor. Domain skipping mutations have been observed both in certain mouse plasmacytomas and in human heavy chain diseases, but deletion of a C_H2 domain is unprecedented in either case (*55*). It would also be surprising that the only two tumors of this class, TEPC 1017 and TEPC 1033, are both afflicted by the same kind of defect. Moreover, the evidence that normal B cells have membrane δ chains similar in size to that of TEPC 1017 suggests that these tumors express a normal form of the δ gene and not an aberrant one.

If mouse in fact has a Cδ2 gene, we would suggest that it may appear only in the secreted form, as seen in the human case. If this is true it would raise the question of control at the level of domain skipping. Specifically, the RNA splicing pattern in the B cell would have to be altered after triggering to include an extra domain in the heavy chain mRNA. Some indirect support for this notion comes from the report by Bargellesi *et al.* (*27*) that serum δ chains of human and mouse appear to be the same size.

Possible Structure-Function Relationships

The unusual structure we determined may be pertinent in elucidating the differential roles of IgD and IgM on B cell membranes. As noted above, the nature of their membrane attachment may be quite different and may further suggest that their heavy chains interact with different molecules on the cell surface. For example, immunofluorescence studies (*56*) suggest that both IgD and IgM are associated with lipopolysaccharide receptor, yet only IgM is associated with the lymphocyte Fc receptor. On the other hand, IgD, but not IgM, appears to associate with the membrane complement C3b receptor (*57*).

A possible correlation of IgD structure with its putative role in triggering B cell differentiation may be provided by the unsual δ chain hinge structure. The extreme susceptibility of membrane IgD to proteolysis under conditions in which IgM remains intact prompted the suggestion (*58*, *59*) that hinge region proteolysis following antigen exposure plays a major role in triggering. Our CδH sequence is not only rich in tryptic-like cleavage sites, but these sites and others may also be more exposed conformationally owing to the absence of inter-δ chain disulfide bridges and scarcity of Pro residues.

These same properties of the δ hinge in conjunction with its somewhat extended length might allow IgD to be cross-linked on the membrane by polyvalent antigens with a higher degree of efficiency than IgM (no hinge region and two inter-μ chain disulfide bridges). There is suggestive evidence (*60*) that membrane IgD is more effective than IgM in triggering B cells by thymus-dependent antigens that have a low density of antigenic determinants.

Conclusion

The data we have obtained provide a look at the primary structure of mouse IgD proteins and their genes. A major portion of the heavy chain mRNA from the IgD plasmacytoma TEPC 1017 and the region of genomic DNA encoding this Cδ mRNA has been cloned and sequenced. The structure we deduced for the plasmacytoma protein is unprecedented and provocative in that only two C_H domains separated by a hinge are found. We have discussed a number of indirect lines of evidence suggesting that this unusual constant region domain structure is characteristic of BALB/c membrane δ chains. Although our data answer some structural questions, they raise many more regarding function. It is tempting to speculate that the exon-shuffling mechanism, as seen for IgM, will provide a flexible and complex array of possible forms of IgD adapted to biological function. The δ gene provides a wealth of potential RNA splice sites that could mediate such functions (*61*).

References and Notes

1. G. Möller, *Immunol. Rev.* **37**, 3 (1977).
2. W. E. Paul, G. W. Siskind, B. Benacerraf, Z. Ovary, *J. Immunol.* **99**, 760 (1967).
3. H. O. McDevitt, Ed., *Ir Genes and Ia Antigens* (Academic Press, New York, 1978).
4. R. K. Gershon and W. E. Paul, *J. Immunol.* **107**, 872 (1967); K. Rajewsky, V. Schirrmacher, S. Nase, N. K. Jerne, *J. Exp. Med.* **129**, 1131 (1969); S. J. Black and L. A. Herzenberg, *ibid.* **150**, 174 (1979).
5. F. S. Walsh and M. J. Crumpton, *Nature (London)* **269**, 307 (1974).
6. E. S. Vitetta and J. W. Uhr, *J. Exp. Med.* **139**, 1599 (1974); U. Melcher, E. S. Vitetta, M. McWilliams, M. E. Lamm, J. M. Phillips-Quagliata, J. W. Uhr, *ibid.* **140**, 1427 (1974).
7. I. ZanBar, S. Strober, E. S. Vitetta, *J. Immunol.* **123**, 925 (1979).
8. L. A. Herzenberg, S. J. Black, T. Tokuhisa, L. A. Herzenberg, *J. Exp. Med.* **151**, 1017 (1980).
9. J. C. Cambier, E. S. Vitetta, J. R. Kettman, G. Wetzel, J. W. Uhr, *ibid.* **146**, 107 (1977); E. S. Vitetta, U. Melcher, M. McWilliams, J. Phillips-Quagliata, M. Lamm, J. W. Uhr, *ibid.* **141**, 206 (1975).
10. R. Okumura, C. M. Metzler, T. T. Tsue, L. A. Herzenberg, L. A. Herzenberg, *ibid.* **144**, 345 (1976).
11. E. S. Vitetta and J. W. Uhr, *Immunol. Rev.* **37**, 506 (1977).
12. P. K. Lala, J. E. Layton, G. J. V. Nossal, *Eur. J. Immunol.* **9**, 39 (1979).
13. J. F. Kearney, M. D. Cooper, J. Klein, E. R. Abney, R. M. E. Parkhouse, A. R. Lawton, *J. Exp. Med.* **146**, 297 (1977).
14. J. W. Goding and J. E. Layton, *ibid.* **144**, 852 (1976).
15. R. L. Coffman and M. Cohn, *J. Immunol.* **118**, 1806 (1977); S. M. Fu, R. J. Winchester, H. G. Kunkel, *ibid.* **114**, 250 (1975).
16. R. J. Poljak, in *CRC Critical Reviews in Biochemistry* **5**, 45 (1978).

17. N. M. Gough, D. J. Kemp, B. M. Tyler, J. M. Adams, S. Cory, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 554 (1980); J. G. Seidman and P. Leder, *Nature (London)* **276**, 790 (1978); C. Brack and S. Tonegawa, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5652 (1977).
18. P. Early, H. Huang, M. Davis, K. Calame, L. Hood, *Cell* **19**, 981 (1980); E. E. Max, J. G. Seidman, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3450 (1979); H. Sakano, K. Hüppi, G. Heinrich, S. Tonegawa, *Nature (London)* **280**, 288 (1979).
19. P. Early, J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, L. Hood, *Cell* **20**, 313 (1980).
20. J. Rogers, P. Early, C. Carter, K. Calame, M. Bond, L. Hood, R. Wall, *ibid.*, p. 303.
21. M. Kehry, C. Sibley, C. Fuhrman, J. Schilling, L. Hood, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2932 (1979); M. Kehry *et al.*, in preparation.
22. V. T. Oi, V. M. Bryan, L. A. Herzenberg, L. A. Herzenberg, *J. Exp. Med.* **151**, 1260 (1980).
23. R. M. E. Parkhouse and M. D. Cooper, *Immunol. Rev.* **37**, 105 (1977).
24. H. L. Spiegelberg, *Nature (London)* **254**, 723 (1975).
25. L.-C. Lin and F. W. Putnam, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6572 (1979).
26. F. D. Finkelman, V. L. Woods, A. Berning, I. Scher, *J. Immunol.* **123**, 1253 (1979); R. Sitia, G. Corte, M. Ferraini, A. Bargellesi, *Eur. J. Immunol.* **7**, 503 (1979); J. Radl, P. VandenBerg, C. M. Jol-Vanderzijde, *J. Immunol.* **124**, 2513 (1979).
27. A. Bargellesi, G. Corte, E. Cosulich, M. Ferrarini. *Eur. J. Immunol.* **9**, 490 (1979).
28. B. Lisowska-Bernstein and P. Vassalli, in *Membrane Receptors of Lymphocytes*, M. Seligmann, J. L. Preud'homme, F. M. Kourilsky, Eds. (North-Holland, Amsterdam, 1975), p. 39; U. Melcher and J. W. Uhr, *J. Immunol.* **116**, 409 (1976); E. R. Abney and R. M. E. Parkhouse, *Nature (London)* **252**, 600 (1974).
29. J. W. Goding and L. A. Herzenberg, *J. Immunol.* **124**, 2540 (1980).
30. J. W. Goding, *ibid.*, p. 2082; S. W. Kessler, V. L. Woods, F. D. Finkelman, I. Sher, *ibid.* **123**, 2772 (1979).
31. L. Eidels, *ibid.* **123**, 896 (1979).
32. M. Mescher and R. R. Pollock, *J. Immunol.* **123**, 1155 (1979); R. R. Pollock, M. E. Dorf, M. F. Mescher, *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 4256 (1980); R. R. Pollock and M. F. Mescher, *J. Immunol.* **124**, 1668 (1980).
33. C.-P. Liu, P. W. Tucker, J. F. Mushinski, F. R. Blattner, *Science* **209**, 1348 (1980).
34. F. D. Finkelman, S. W. Kessler, J. F. Mushinski, M. Potter, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **39**, 481 (1980); S. W. Kessler, J. F. Mushinski, M. Potter, F. D. Finkelman, *ibid.*, p. 1055.
35. V. T. Oi, P. P. Jones, J. W. Goding, L. A. Herzenberg, L. A. Herzenberg, *Curr. Top. Microbiol. Immunol.* **81**, 115 (1978).
36. J. F. Mushinski, F. R. Blattner, J. D. Owens, F. D. Finkelman, S. W. Kessler, L. Fitzmaurice, M. Potter, P. W. Tucker, in preparation.
37. T. Maniatis, S. G. Kee, A. Efstratiadis, F. C. Kafatos, *Cell* **8**, 163 (1976).
38. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
39. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
40. Abbreviations: Ala, alanine; Cys, cysteine; Asp, aspartic acid; Glu, glutamic acid; Phe, phenylalanine; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Asn, asparagine; Pro, proline; Gln, glutamine; Arg, arginine; Ser, serine; Thr, threonine; Val, valine; Trp, tryptophan; and Tyr, tyrosine.
41. E. A. Kabat, T. T. Wu, H. Bilofsky, Eds., *Sequences of Immunoglobulin Chains* (NIH Publ. 80-2008, 1979), p. 138.
42. M. R. Lerner, J. A. Boyle, S. M. Mount, S. L. Wolin, J. A. Steitz, *Nature (London)* **283**, 220 (1980); J. Rogers and R. Wall, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1877 (1980).
43. T. Honjo, M. Obata, Y. Yamawaki-Kataoku, Y. Kataoku, T. Kawakami, H. Takahashi, Y. Mano, *Cell* **18**, 559 (1979); P. W. Tucker, K. Marcu, N. Newell, J. R. Richards, F. R. Blattner, *Science* **206**, 1303 (1979).
44. L.-C. Lin and F. W. Putnam, personal communication.
45. M. Tomita and V. T. Marchesi, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2964 (1975).
46. D. D. Jones, *J. Theor. Biol.* **50**, 167 (1975).
47. G. Guidotti, *J. Supramol. Struct.* **7**, 489 (1977); C. Tanford and J. A. Reynolds, *Biochim. Biophys. Acta* **457**, 133 (1976).
48. J. P. Segrest and R. J. Feldmann, *J. Mol. Biol.* **81**, 853 (1974).
49. P. W. Robbins, S. C. Hubbard, S. J. Turco, D. F. Wirth, *Cell* **12**, 893 (1977).
50. R. Breathnach, C. Benoist, K. O'Hare, F. Gannon, P. Chambon, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4853 (1978).
51. N. J. Proudfoot and G. G. Brownlee, *Nature (London)* **263**, 211 (1976).
52. S. Rudikoff, personal communication.
53. R. L. Jilka and S. Pestka, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5692 (1977).
54. S. W. Kessler and F. D. Finkelman, personal communication.
55. E. C. Franklin and B. Frangioni, in *Contemporary Topics in Molecular Immunology*, F. P. Inman and W. J. Mandy, Eds. (Plenum, New York, 1976), p. 89; E. C. Franklin, personal communication.
56. L. Forni and A. Coutinho, *Nature (London)* **273**, 304 (1978).
57. R. Sitia, E. M. Rabellino, M. Sockell, U. Hammerling, *J. Immunol.*, in press.
58. E. S. Vitetta and J. W. Uhr, *Science* **189**, 964 (1975).
59. A. Bourgois, E. R. Abney, R. M. E. Parkhouse, *Eur. J. Immunol.* **7**, 210 (1977).
60. E. Pure and E. S. Vitetta, *J. Immunol.* **125**, 420 (1980).
61. *Note added in proof*: Partial amino acid sequence data on proteolytic fragments derived from a shortened δ chain secreted by rat IgD myeloma were recently reported [G. Alcaraz, A. Bourgois, A. Moulin, H. Bazin, M. Fougereau, *Ann. Immunol. (Inst. Pasteur)* **131C**, 363 (1980)]. Two fragments can be matched with high homology to corresponding regions in our mouse sequence, and the authors postulated that the shortened length of the rat δ chain resulted from a partial deletion of the Cδ2 domain. However, this deduction was based, at least in part, on erroneous homology alignment, and thus conclusive cross-species comparison must await additional refinement of the rat sequence.
62. K. W. Olsen, *Biochim. Biophys. Acta* **622**, 259 (1980).
63. H. B. Bull and K. Breese, *Arch. Biochem. Biophys.* **161**, 665 (1974).
64. Y. Nozaki and C. Tanford, *J. Biol. Chem.* **246**, 2211 (1971).
65. P. Y. Chou and G. B. Fasman, *Biochem.* **13**, 211 (1974).
66. We thank M. Fiant for electron microscopy, H.-L. Cheng and R. Robinson for expert technical assistance, W. Fiske and E. LaLuzene for help with the manuscript, S. Rudikoff for amino acid sequence determinations, L. Fitzmaurice, V. Oi, K. Olsen, J. Schroeder, J. Uhr, and F. Finkelman for helpful discussions, and M. Potter for enthusiastic encouragement and for providing the plasmacytomas that made this work possible. This work was done under the applicable NIH recombinant DNA guidelines and was supported by NIH grants AI 16547 (P.W.T.), GM 21812 (F.R.B.), and GM 06768 (C.P.L.).

16 July 1980

DNA Sequences Mediating Class Switching in α-Immunoglobulins

Mark M. Davis, Stuart K. Kim, Leroy E. Hood

The antibody molecule is a set of discrete molecular domains that carry out two general types of functions. The variable (V) domain binds antigen and the constant (C) domains trigger effector functions such as complement fixation. The V and C domains arise from the interactions of two different polypeptides, light (L) and heavy (H), which in turn are encoded by a series of discrete gene segments—V_L, J_L (joining), and C_L encoding the light chains and V_H, D (diversity), J_H, and C_H encoding the heavy chains (*1–4*). During the differentiation of antibody-producing or B cells, two distinct types of DNA rearrangements of these gene segments occur (*4, 5*). One type generates the V_L gene by direct joining of the V_L and J_L gene segments and the V_H gene by direct joining of the V_H, D, and J_H gene segments. These DNA rearrangements are termed V-J or V-D-J joining and they are, in part, responsible for the generation of antigen-binding diversity in V domains.

A second type of DNA rearrangement, termed C_H switching, allows important flexibility in the use of a given antigen-binding site. At an early stage of B cell differentiation, an individual B cell initially expresses immunoglobulin M (IgM) molecules with a single V domain (V_L-V_H combination) (*6, 7*). Later, this B cell or its clonal progeny may express another immunoglobulin class while continuing to express the same V domain (*8*). Since the class of immunoglobulin is determined by the C_H region (C_μ, C_γ, and C_α determining IgM, IgG, and IgA, respectively), the B cell must shift from the expression of another C_μ gene to the expression of another C_H gene during differentiation. Thus, C_H switching associates a particular antigen-binding specificity, the V domain, with a series of different effector functions encoded by the various C_H regions.

Two types of experiments have provided insights into the mechanism of C_H switching. First, Honjo and Kataoka (*9*)

M. M. Davis and S. K. Kim are graduate students and L. E. Hood is professor of biology at the Division of Biology, California Institute of Technology, Pasadena 91125.