GALACTAN-BINDING ANTIBODIES

Diversity and Structure of Idiotypes

By STUART RUDIKOFF, MICHAEL PAWLITA, JANET PUMPHREY, ELIZABETH MUSHINSKI, AND MICHAEL POTTER

From the Laboratory of Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Two of the intriguing questions in immunology, which in many instances may be related, are the mechanisms by which the vast array of immunoglobulin diversity is generated and the structure of antigenic determinants associated with these molecules. It is now clear that several processes contribute to immunoglobulin diversity. First, there are an apparently large number of both light (L)[1] and heavy (H) chain germline genes (1–4). Second, both L (5–8) and H (9–18) polypeptide chains are encoded in multiple genetic elements, and various combinations of these segments produce a large amount of structural diversity. Third, the joining of these various gene segments may not be precise (6–8, 11–13, 17–21), resulting in diversity at the points of recombination. Fourth, L and H chains may randomly pair, although the actual extent of such chain 'shuffling' is undetermined. Fifth, somatic point mutations occur that impose additional diversity on the system (22–29). Sixth, gene interaction (such as gene conversion) among related members of immunoglobulin variable (V) region families may alter primary sequences (30–32). It is thus of interest to analyze the primary structure of groups of related antibodies in terms of the genetic mechanisms that have contributed to the generation of individual molecules. Comparison of related antibodies further presents the opportunity to assess the structural basis of antigenic determinants associated with these molecules and analyze the role of the various genetic processes in the generation of these markers.

Since the original discovery in humans (33) and rabbits (34) of individual antigenic specificities (subsequently termed idiotypes) associated with homogeneous immunoglobulins, the structure and potential biologic role of these markers have been of great interest in the field of immunology. Numerous idiotypic systems have been generated and studied in great detail in a variety of species. Most of these studies have dealt with serological characterization of the antibodies involved and very little is known about the molecular basis of idiotypy. Interest in the potential role of idiotypes as biological regulators has been stimulated by the proposal of Jerne (35) suggesting that these determinants constitute a network that may serve to regulate the immune response. In view of the potential value of idiotypes as models for protein antigens, as well as their possible role in immune regulation, it is important to understand the molecular basis of these determinants. One of the systems being used in this laboratory to assess this

[1] *Abbreviations used in this paper:* L, light chain; H, heavy chain; V, variable region.

question consists of myeloma and hybridoma antibodies directed to β(1,6)-D-galactan moieties. In the present communication we report idiotypic serological analysis and complete H chain V region sequences from eight of these proteins. These results have permitted the subsequent assignment of the molecular basis of a number of idiotypes associated with galactan-binding antibodies, as well as provided insights into the genetic mechanisms contributing to the structure of these proteins.

Materials and Methods

Antisera and Serological Assays. Antiidiotypic antisera were prepared in A/He and SJL/J mice and tested in a solid-phase radioimmunoassay as previously described (36). Briefly, antisera were first assayed for direct binding of ^{125}I-labeled immunogen. The highest dilution giving reproducible binding (usually ~2,000 cpm) was then used in a competition assay using various 'cold' hybridoma and myeloma proteins to compete with labeled immunogen.

Proteins. The production and characterization of hybridoma proteins with specificity for β(1,6)-D-galactan has previously been reported (37). These hybridomas were derived from three separate fusions. Hybridoma proteins included in the present studies were of the IgM class and myeloma proteins were of the IgA class. All antibodies were purified by affinity chromatography (38).

Peptides and Sequence Determination. H chains were cleaved with cyanogen bromide (19) and the resulting fragments separated on a Sephadex G-100 column equilibrated in 5 M guanidine–0.2 M NH_4HCO_3. Peptides linked by disulfide bonds were completely reduced, ^{14}C-alkylated, and resolved on a Bio-Gel A, 0.5 M agarose column equilibrated in 6 M guanidinium-Tris buffer, pH 8.0. Sequence determinations were performed on a modified Beckman 890C sequencer (Beckman Instruments, Inc., Fullerton, CA) (30) and the phenylthiohydantoin derivatives obtained after each degradation cycle were identified by high pressure liquid chromatography (39).

Results

Sequence Determination. H chain V region peptides were isolated and sequenced as previously described (30). Complete V region sequences from eight H chains are presented in Fig. 1 along with sequences of myeloma proteins previously reported (10). Four of the hybridoma H chains, HyGal 1, 2, 3, and 4, are identical to each other and to the myeloma proteins X44 and T601 for amino acids 1–94, the region encoded by the V_H gene segment. Proteins HyGal 6, 11, and 12 share the same Asp-Glu substitution at position 46. HyGal 6 has two additional substitutions at residues 45 and 91. HyGal 10 differs from the prototype sequence at positions 60, 88, and 91. The D regions of all hybridoma proteins except HyGal 11 and 12 differ and the D region of HyGal 3 is identical to those of T601 and X24. At least three J_H regions are expressed among these proteins (Fig. 1) and the HyGal 11 and 12 J_H regions have the identical Ala-Asp substitution at position 101.

Idiotypic Analysis. Individual allogeneic antisera raised to each of the hybridoma and myeloma proteins were tested in competition assays using each of the other galactan-binding proteins as inhibitors (Table I). Antisera to hybridoma proteins HyGal 1, 2, and 4 and myeloma proteins X24, X44, and J539 were considered to be specific for the immunogen and showed essentially no cross-reactivity. In contrast, antisera to hybridoma HyGal 3 cross-reacted with proteins HyGal 2, T601, and X24. Antisera to HyGal 6 reacted equally well with HyGal

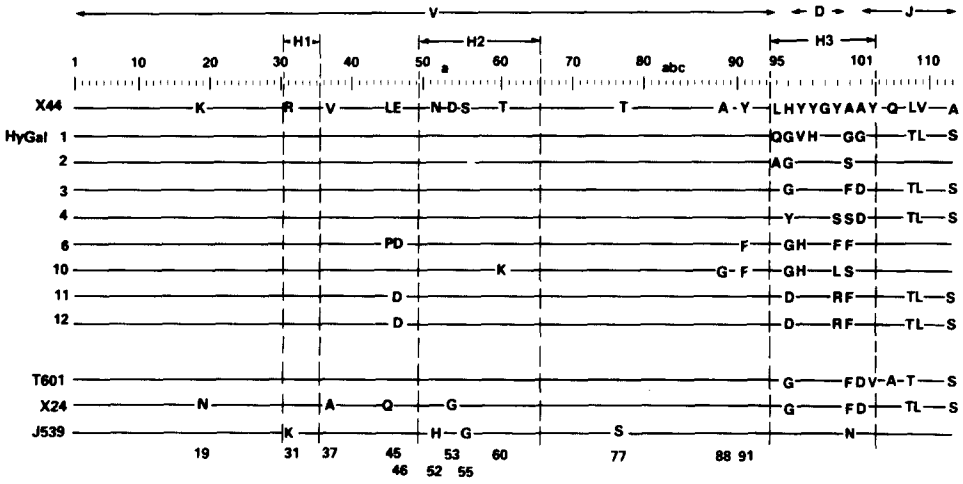


FIGURE 1. Complete H chain V region sequences from BALB/c galactan-binding hybridoma proteins. Sequences for the myeloma proteins X44, T601, X24, and J539 are from reference 10. Numbers at the bottom indicate positions at which substitutions occur compared with the X44 sequence, which is exactly encoded by a BALB/c V_H gene (42). Numbering is according to Kabat et al. (57). Gln-81 in HG3 and Thr-44 in HG11 are tentative assignments.

TABLE I

*Idiotype Expression by Galactan-binding Antibodies**

Inhibitors	Idiotypic systems‡										
	HG1	HG2	HG3	HG4	HG6	HG10	HG12	T601	X44	X24	J539
Hybridomas											
HG1	[13]	>	>	>	>	>	>	>	>	>	>
HG2	>	[11]	[100]	>	>	>	[100]	>	>	>	>
HG3	300	>	[5]	>	>	>	[20]	[35]	>	>	>
HG4	>	>	700	[50]	>	>	[30]	>	>	>	>
HG6	>	>	>	>	[10]	[10]	[20]	>	>	>	>
HG10	>	>	>	>	[12]	[9]	[20]	>	>	>	>
HG12	>	>	250	>	>	500	[20]	>	>	>	>
Myelomas											
T601	>	>	[5]	>	>	>	>	[15]	>	>	>
X44	>	>	>	>	>	>	>	[150]	[6]	>	>
X24	>	>	[10]	>	>	>	>	[50]	>	[10]	>
J539	>	>	>	>	>	>	>	>	>	>	[3]

* Antisera to specified hybridoma or myeloma proteins were tested for competition of binding of radiolabeled immunogen. Values given represent nanograms required for 50% inhibition. > indicates 50% inhibition was not achieved at concentrations of 1,000 ng or greater. Enclosed values represent significant inhibition arbitrarily defined as requiring <20 times the amount of homologous protein.

‡ Idiotypic antisera were prepared to the hybridoma and myeloma proteins indicated and tested with inhibitors listed in column at left.

10 and the reciprocal activity was found with antisera to HyGal 10. Antisera to HyGal 12 showed broader specificity than observed in other systems and reacted with all IgM hybridoma proteins except HyGal 1, but with none of the IgA myeloma proteins. Antisera to T601 cross-reacted significantly with HyGal 3, T601, X24, and X44.

Discussion

The study of groups of myeloma and hybridoma proteins with specificity for chemically defined haptenic determinants has provided a number of systems that have proven to be of great value in the analysis of antibody diversity and the structure of idiotypes. One such system being developed in our laboratory consists of antibodies with specificity for β(1,6)-D-galactan moieties. A number of myeloma proteins demonstrating this specificity were initially characterized in terms of specificity (38, 40), idiotypy (41), and primary sequence analysis (10, 19). More recently we have generated a series of hybridoma proteins exhibiting the same specificity and have reported complete κ chain V region sequences for 10 of these (37). In the present study we have determined complete H chain V region sequences for eight of the hybridoma proteins. In addition, idiotypic antisera have been raised to each of the hybridoma and myeloma proteins and a detailed serological analysis has been performed. The evaluation of this data has provided information pertinent to both the generation of antibody diversity and the structure of immunoglobulin idiotypes.

V_H Segment Sequence Diversity. The H chain V region is encoded in three distinct genetic elements (12, 13) designated V_H, D, and J_H. The V_H gene segment encodes amino acids 1~95, the D segment encodes a portion of the third complementarity determining region, and the J_H segment encodes amino acids 101–113 (Fig. 1). A comparison of the V_H sequences in Fig. 1 reveals that hybridoma proteins HyGal 1, 2, 3, and 4 and myeloma proteins X44 and T601 are identical for their first 94 amino acids. The V_H segment in these proteins is likely to be 94 amino acids in length, since Ollo et al. (42) have cloned and sequenced a V_H germline gene that encodes exactly this sequence. HyGal 6 differs from this sequence at three positions (45, 46, 91), HyGal 10 at three positions (60, 88, 91), HyGal 11 at one position (46), and HyGal 12 at one position (46). The Asp-Glu substitution at position 46 is shared by proteins HyGal 6, 11, and 12. All of these substitutions can be accounted for by single-base nucleic acid changes.

It is of interest to now ask, what is the origin of the V_H sequences found in the antigalactan hybridomas? The HyGal 1, 2, 3, 4, X44 and T601 sequences can be directly encoded by the V_H germline gene sequenced by Ollo et al. (42). Our laboratory has prepared a cDNA probe to the X24 V_H region with which, in Southern blot analysis of genomic DNA (A. Hartman and S. Rudikoff, unpublished results), we detected two hybridizing bands. Thus, the V_H regions expressed in these proteins are likely to be encoded by one or at most two genes. If, for example, a second gene encoded the V_H segments with Asp at position 46 (HyGal 6, 11, 12), there would still exist a number of substitutions apparently not encoded in germline genes, i.e., Pro-45 in HyGal 6, Lys-60 in HyGal 10, Gly 88 in HyGal 10, and possibly Phe-91 in HyGal 6 and 10. These substitutions

would, presumably, have arisen by somatic mutation. Proteins HyGal 11 and12 are identical in sequence in both their L and H chains and originated from the same fusion. These hybridomas are therefore probably derivatives of the same original clone so that Asp-46 has occurred independently only twice (once in HyGal 6 and once in the HyGal 11, 12 pair). HyGal 6 and 10 are similarly notable in that they arose from the same fusion, have identical J_κ and J_H segments, and share the Phe-91 substitution. It is thus possible that these hybridomas derive from a single precursor that expressed the Phe-91 substitution and have subsequently diversified by additional somatic mutations as observed in Fig. 1.

The role and time of occurrence of somatic mutation in immunoglobulin diversity has long been a subject of controversy. The occurrence of somatic mutation in immunoglobulin genes has been well documented (22–29) and more recently it has been suggested (25–27) that the somatic mutation process is linked to Ig class switching. The fact that the galactan-binding IgA myeloma proteins have a number of substitutions likely resulting from somatic mutation is consistent with this interpretation. However, the putative somatic mutations observed in the IgM hybridoma proteins are not associated with class switching. We interpret the present data to indicate that somatic mutation does occur in IgM antibodies. It has been suggested (37, 43) that the failure to previously detect this type of event in IgM-producing cells is related to the observation that IgM-producing cells are not derived from a memory population. Thus, IgM-producing cells, including those expressing somatic mutations, would be rapidly turned over and lost from the population. Lymphocytes expressing other H chain classes do arise from a memory pool and might thus have a higher probability of exhibiting somatic mutations if this process is correlated with time. At present, there is no direct evidence linking the occurrence of somatic mutation to class switching although this possibility certainly exists.

Complementarity-determining Region 3 and D Segments. An examination of Fig. 1 reveals that all of the hybridoma proteins are identical in CDR-1 and -2 (H1, H2) with the exception of a single substitution in CDR-2 of HyGal 10. However, CDR-3 (H3) shows a high degree of sequence variation. This region of the molecule is of particular interest in that a portion of CDR-3 is encoded by the D gene and this region is also the site of two recombination events, V_H-D and D-J_H.

We have analyzed the CDR-3 sequences in terms of their origin and this data is presented in Fig. 2. Numbering in the figure differs from that in Fig. 1 and has been adapted to best present the derivation of amino acids found in CDR-3. The consensus protein sequence found in Fig. 1 is presented at the top of the figure. The end of the V_H segment has been assigned to position 94 based on the V_H gene sequence determined by Ollo et al. (42) and the J_H sequences at the right-hand side are those of the three J_H segments used in the galactan-binding antibodies. It can be seen that the 'core' sequence of CDR-3, designated D, can be encoded by at least six different germline D genes (18) listed in the lower portion of the figure. However, it is readily apparent that a number of amino acids are present on both the NH_2- and carboxy-terminal sides of the D segment (designated *x* and *y*, respectively) which are not encoded in these or any other known germline D segments, and are presumably generated during the V_H-D and D-J_H recombination events (see below).

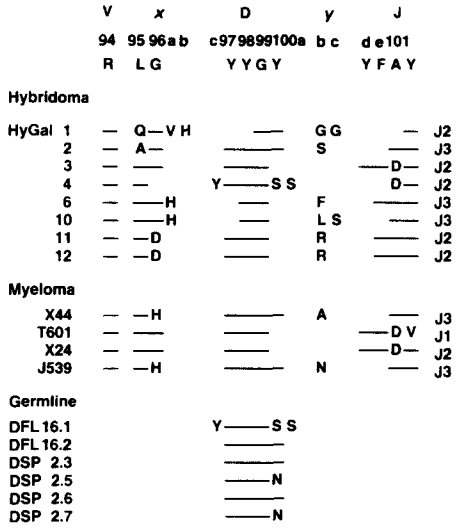


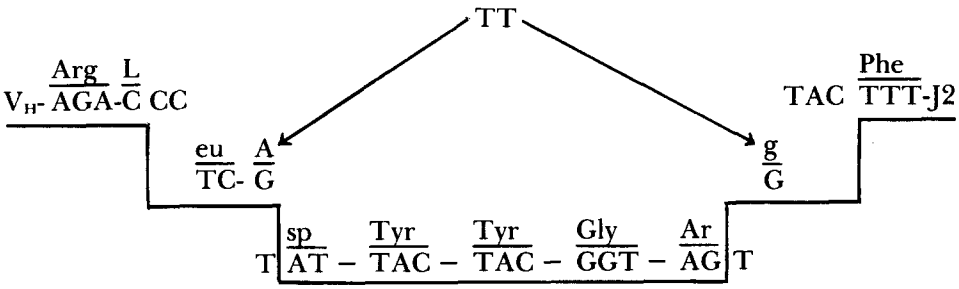
FIGURE 2. H chain CDR-3 sequences found in galactan-binding hybridoma and myeloma proteins. The *x* and *y* regions designate amino acids on the NH_2 and COOH sides of the D segment that are not encoded by germline D genes. The prototype sequence R-L-G-Y-Y-G-Y-Y-F-A-Y represents the most frequent amino acid found at the designated position although this exact sequence is not expressed in any of the hybridoma or myeloma proteins. The germline D sequences are translates of the corresponding germline genes.

The derivation of several of the D region sequences appears relatively unambiguous. The HyGal 4 D region is directly encoded by the DFL 16.1 gene. This D gene probably also encodes the HyGal 11 and 12 D regions. Three of the amino acids in the HyGal 11 D region are directly encoded by DFL 16.1. The germline codon for Tyr-96C is TAT and an alteration of this codon to GAT during V_H-D recombination would produce the Asp-96 observed in HyGal 11 and 12. The D regions of HyGal 1, 2, 3, 6, 10, T601, and X24 could be encoded by several of the listed D region genes. The X44 and J539 D segments may be encoded by the DFL 16.2 gene since translation of the trinucleotide 5′ to that encoding Tyr-97 would produce the His observed at position 96. The portion of CDR-3 encoded by the D gene appears to vary from 2 (HyGal 1) to 6 amino acids although, as will be seen below, the recombination events compensate to maintain a uniform length in CDR-3.

Probably the most striking feature of Fig. 2 is the number of amino acids NH_2- and carboxy-terminal (*x* and *y*) to those encoded by germline D segments. None of these sequences appear to be encoded in D genes and are unlikely to have been generated by D-D fusion. It has previously been suggested (44, 45) that nucleotides encoding amino acids on the 5′ and 3′ sides of D segments are generated by repair enzymes such as terminal transferase. Most of the amino acids in the *x* and *y* regions of the galactan-binding H chains are high in G-C content, which is consistent with the proposed use of terminal transferase as a means for introducing amino acids in these areas. An example of the potential role of repair enzymes in the generation of sequences in the *x* and *y* regions is presented for the HyGal 12 sequence. In HyGal 12, which we assume uses DFL

16.1, it is necessary to generate Leu-95, Asp-96 during V_H-D recombination and Arg-100b during D-J_H recombination. A possible scheme for these events is as follows:

DFL 16.1



The first nucleotide of Leu-95 would be contributed by the codon 3′ to the V_H gene (CCC). The remaining two nucleotides (TC) and the first nucleotide (G) of Asp-96 would be added by terminal transferase (TT), followed by recombination with DFL 16.1 as indicated. On the 3′ side the first two nucleotides of the AGT (Ser) codon would be used, followed by an addition of G by terminal transferase to generate Arg 100b (AGG) and recombination with J2. Similar schemes can be devised to account for most of the sequences in the *x* and *y* regions.

D Region Function. Examination of the CDR-3 and D sequences (Figs. 1 and 2) reveals that only Gly-99 is invariant and that a large amount of amino acid sequence variation is present in this region of the molecule. It should be noted that all hybridomas were induced by immunization with β(1,6)-D-galactan containing antigens and were hapten eluted from affinity columns. Therefore, the variation in CDR-3 contributed by D region sequence and the two joining events produces no significant alteration in specificity. Furthermore, the L chains used in these antibodies are essentially identical (37), indicating that these CDR-3 and D region sequences do not effect H and L pairing. The only constant feature of this region of the molecule is the invariant length of CDR-3. The length of CDR-3 may thus be important in either the general architecture of the binding site or the ability of these H chains to pair with a proper L chain, but the primary amino acid sequence seems relatively unimportant. We have made similar observations for phosphocholine and dextran-binding antibodies (45, 46). It is somewhat surprising to find that in the first three systems studied in appropriate detail (phosphocholine, dextran, and galactan), D region sequence appears to play no important role in either the determination of antigen-binding specificity or in H-L pairing.

Alternate Frames of Recombination. In spite of the variation in length of the D segment (two to six amino acids) the length of CDR-3 invariably consists of nine amino acids between positions 95 and 101 (Figs. 1 and 2). The variation in length of the D segment is compensated by alterations in the frame of V_H-D recombination on the 5′ side of D and D-J_H recombination on the 3′ side to maintain the length of CDR-3. Examples of this alteration in recombination sites between D and J are presented in Fig. 3. By similar alterations in the V_H-D recombination

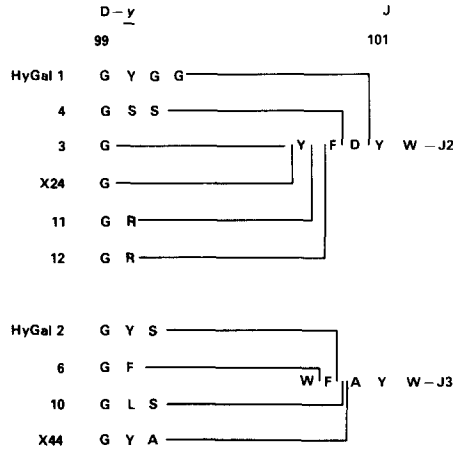


FIGURE 3. Alternate sites of D-J_H recombination in galactan-binding antibodies.

sites, the length of CDR-3 is maintained, implying that this conserved length is probably critical to some functional aspect of these H chains.

Idiotypes. Idiotypic markers, which may be defined generally as antigenic determinants within the immunoglobulin V region, have proven to be extremely valuable tools in the areas of immunogenetics and structure-function relationships. More recently, interest in idiotypes has expanded with the hypothesis of Jerne (35) suggesting these markers may be recognition sites for the regulation of immune responses. Unlike the constant region allotypes, which in many instances are simple in nature and determined by one or a few amino acids (47), idiotypes appear to be quite complex and have been difficult to characterize structurally (reviewed in 46).

Idiotypes of Galactan-binding Hybridomas Are Located in the H Chain CDR-3 Region. We have approached the problem of the structural basis of idiotypes by isolating a group of antibodies reacting with the same haptenic determinant and performing complete L and H chain V region sequence analysis in conjunction with serologic antiidiotypic studies. The combined results of these studies now permit an assessment of the structural basis of idiotypes expressed on these molecules.

We have previously reported (37) complete L chain V region sequences for all proteins included in the present study. The L chains from seven of eight hybridomas and two myelomas, X24 and T601, are identical throughout their V_κ regions but use different J_κ segments. The HyGal 6 L chain has a single Ala-Val substitution at position 12. However, this substitution is not involved in an idiotypic determinant, as antisera to HyGal 6 react equally well with HyGal 10 (Table I), which expresses the 'germline' Ala at this residue. Similarly, no common reaction pattern is observed among proteins sharing the same J_κ sequence, indicating J_κ is not a primary source of idiotypic determinants. Two of the myeloma L chains, J539 and X44, have V_κ substitutions that may contribute to idiotypic markers. The structural basis of idiotypes associated with all hybridoma as well as myeloma proteins X24 and T601 must therefore reside in the H chains. Previous experiments using recombinant molecules made among the

myeloma proteins have likewise shown that idiotypes expressed on these molecules are determined by the H chain sequence (48).

Examination of the H chain sequences (Fig. 1) reveals very few substitutions in the V_H segments (amino acids 1–94). Three proteins, HyGal 6, 11, and 12 share the same Asp-46 substitution. However, antisera to HyGal 6 do not cross-react with HyGal 12 and antisera to HyGal 12 cross-react with HyGal 6 as well as proteins with glutamic acid at residue 46, indicating Asp-46 is not an idiotype-determining residue. Similarly, no common reaction pattern is observed among proteins with Glu at position 46. Since HyGal 6 and 10 show strong cross-reactivity, the Pro-45 substitution in HyGal 6 and the Lys-60 and Gly-88 substitutions in HyGal 10, by the same argument, are not the basis of idiotypic determinants. The shared Tyr-Phe substitution at residue 91 in these two proteins is also not a likely source of idiotypic determinants (see below). The lack of cross-reactions among proteins sharing the same J_H segment also eliminates this region as a primary idiotypic determinant. Two of the myeloma proteins, X24 and J539, have multiple V_H substitutions that are all potentially available for involvement in antigenic markers. Thus, idiotype-determining residues in all hybridoma proteins, as well as myeloma proteins X44 and T601, reside in the CDR-3 region and result either from (*a*) the D segment, (*b*) the two recombination events occurring in this region, or (*c*) combinations of the above three elements.

Structural Basis of Antigalactan Idiotypes. Since most idiotypes expressed on the galactan-binding antibodies are localized in CDR-3, an examination of the sequences in this area (Fig. 4) along with the serologic analysis (Table I) permits a definition of the structural basis of many of these markers. It should be kept in mind that these markers may be complex in nature and involve three-dimensional interactions presently unknown, including invariant portions of the

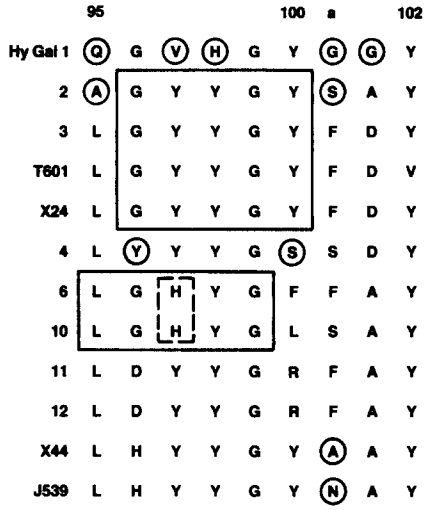


FIGURE 4. Correlation of idiotype expression and H chain CDR-3 sequences. Circled or boxed amino acids represent residues most likely involved in the determination of idiotypes as described in the text.

molecule. In this regard, a hypothetical three-dimensional model of a galactan-binding antibody has been constructed along lines previously described (49), which, within the limits of model building, suggests that the CDR-3 region is exposed to solvent and positioned such that it might be expected to be accessible to antiidiotypic antibodies (Fig. 5).

HyGal 1. Idiotypic antisera to HyGal 1 (Table I) were essentially specific for the immunogen although weak cross-reactivity with HyGal 3 was observed. The CDR-3 sequence of HyGal 1 (Fig. 4) is unique, as might be expected from the serological analysis, and displays little homology with the other proteins.

HyGal 2. Antisera to HyGal 2 show no cross-reactivity with other proteins in the group. However, HyGal 2 is highly homologous to HyGal 3, T601, and X24 with differences occurring only at positions 95, 100a, and 101. This result indicates that Ala-95 or Ser-100a are dominant determinants which alter or mask the central sequence (amino acids 96–100) such that antisera to the HyGal 2 immunogen are non–cross-reactive. It is more likely that these residues are dominant determinants in that antisera to HyGal 3 also cross-react with HyGal 2.

HyGal 3. Antisera to HyGal 3 cross-react with HyGal 2, T601, and X24. Proteins HyGal 3, T601, and X24 have identical sequences from positions 95 to 101 and HyGal 2 differs only at positions 95, 100a, and 101. The strong cross-reactivity between HyGal 2 and 3 suggests that a significant portion of antiidiotypic antibody in this sera is directed to the sequence including positions 96–

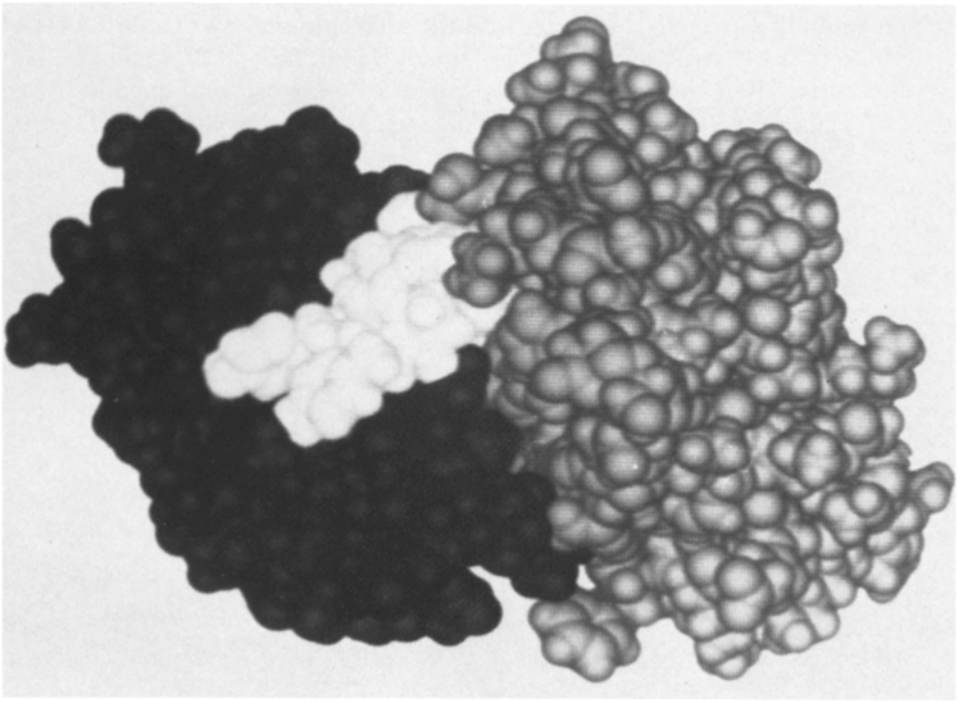


FIGURE 5. Hypothetical three-dimensional model of the galactan-binding myeloma protein J539. Colors are as follows: L chain, black; H chain, grey; H chain CDR-3 region, white.

100, which makes this group unique when compared with the other members. Proteins HyGal 4 and 12, which share a portion of this sequence, were found to weakly inhibit. The complexity, as well as the precise requirements of determinants such as these, is pointed out by the failure of X44 and J539 to inhibit. Based on sequence comparisons, these proteins might well be expected to inhibit this antisera as well as HyGal 4 and 12.

Antisera to T601 react with HyGal 3 and X24 to approximately the same extent, but do not react with HyGal 2, suggesting that either or both positions 95 and 100a are important in the generation of this determinant. Neither Leu-95 nor Phe-100a alone can be dominant residues since other proteins expressing these amino acids do not inhibit. These residues must therefore function in association with adjacent amino acids to produce the cross-reactive determinant recognized in this system. It seems possible that Phe-100a and Asp-101 are the more important idiotype-determining amino acids since this sequence is unique to proteins HyGal 3, X24, and T601. X44 also inhibits this system although the basis for this reactivity is not obvious and X44 lacks Phe-100a and Asp-101, which correlate with idiotype expression in the other molecules. The variation in reactivity patterns between antisera to HyGal 3 and T601 points out the fact that individual antisera may recognize the same protein sequence differently.

HyGal 4. Antisera to HyGal 4 were specific for the immunogen. The HyGal 4 CDR-3 sequence has unique substitutions at positions 96 (Tyr) and 100 (Ser). Either or both of these residues are therefore idiotype determining.

HyGal 6 and 10. Antisera to HyGal 6 react equally well with HyGal 10 and the reciprocal reactivity is seen with antisera to HyGal 10. These antisera show no significant additional cross-reactions with other members of the group. The CDR-3 protein sequences from HyGal 6 and 10 are identical from positions 95 to 99 and both express the unique His-97 substitution, which distinguishes these two proteins in this region. Additionally these proteins share the Phe-91 substitution (Fig. 1) in the framework region adjacent to CDR-3. However, the three-dimensional structure of the galactan-binding myeloma protein J539 (50; D. R. Davies, personal communication), indicates that position 91 is buried in the L-H interface and would not be accessible as an idiotype-determining residue. Thus, His-97 is likely to be the dominant amino acid in the determination of this shared idiotype.

HyGal 12. Antisera to HyGal 12 cross-react broadly with most of the hybridoma proteins, yet do not react with other IgM molecules, indicating that this reactivity is not allotypic. The hybridoma proteins inhibiting this system share the common sequence Tyr-98, Gly-99, yet myeloma proteins of the IgA class with an identical sequence fail to inhibit. This result suggests that the C_H1 domain may effect the expression of idiotypes recognized by this antisera. Similar results have been reported for a monoclonal antibody to phosphocholine (51).

X44. Antisera to X44 are inhibited only by the immunogen, and this CDR-3 sequence is distinguished by the amino acids His-96 and Ala-100a. J539 shares the same His-96 substitution (but differs at 100a) and does not inhibit this system. Additionally, the X44 L chain differs from the other L chains in this group by a Trp-Ile interchange at position 96, the point of V_κ-J_κ recombination. However, position L-96 is buried in the interior of the molecule and would not contribute

to surface determinants. The predominant idiotype-determining residue is therefore Ala-100a in the H chain.

X24. A/J antisera to X24 are specific to the immunogen, although this protein is inhibitory in both the HyGal 3 and T601 systems described above. Thus, antisera to X24 do not appear to be directed to the CDR-3 sequence shared with proteins T601 and HyGal 3. Although X24 has a number of V_H substitutions (Fig. 1), we have previously concluded that the idiotype recognized by this antisera is likely to be determined by Gly-53 (52). In contrast, a monoclonal antibody to X24 was found to cross-react moderately with X44 and weakly with several other of the galactan-binding myeloma proteins (36).

J539. Antisera to J539 were unique for the immunogen. J539 expresses a number of substitutions throughout the H chain V region as well as the V_κ segment, making it impossible to assign idiotypic correlates.

The above analysis has provided the opportunity to describe at the molecular level the likely basis for a number of idiotypes expressed by a group of closely related proteins. Some of these idiotypes, such as those identified by the HyGal 3 and HyGal 6 systems, are shared while others are unique. The CDR-3 sequences of these proteins (Fig. 4) point out the dramatic effect of relatively few amino acid changes on the expression of idiotypes. For example, the idiotype-determining residues in the cross-reactive HyGal 3 system largely include the central sequence of CDR-3. Yet proteins HyGal 12, X44, and J539, which differ by only two amino acids each in this region, either fail to inhibit or at best inhibit weakly. Thus, this small number of substitutions has resulted in a loss or significant alteration of idiotype expression accompanied by the generation of new determinants. This is further illustrated by proteins HyGal 6 and 10, which are distinguished primarily by a single substitution at residue 97. Antisera to either of these proteins are inhibited to the same extent by both, yet no other proteins react in these systems. Thus, proteins that are nearly 99% homologous in their sequences may be completely different idiotypically depending on the particular reagent used. Taken together, these results point out the complex interactions of amino acids in CDR-3 that generate the observed idiotypic determinants and the very pronounced effect resulting from perturbation of these structures.

The study of the structural basis of idiotypy is still in its early stages (reviewed in 46) and very few systems to date have been amenable to rigorous analysis. Two L chain idiotypes have been defined on inulin-binding myeloma proteins (53, 54). The most comprehensive data to date on idiotype structure has been derived from the study of antibodies that bind α(1,3) dextran (55, 56). Two classes of idiotypes have been described in this system. The first, a cross-reacting idiotype located in CDR-2, and the second, a series of individual idiotypes located in the CDR-3 region, as is the case for the galactan-binding antibodies described above. It should be noted that the assignment of the dextran markers assumes that all of the L chains (which are λ type) are likely to be identical although this has not been proven experimentally. Thus, in both the galactan and dextran systems it is possible to define at the molecular level the structural basis for a number of both cross-reacting and individual idiotypes.

Possible Functional Role of Idiotypes. It is interesting that in the galactan system apparently all of the idiotypic determinants recognized by homologous antisera

reside in the most variable part of the immunoglobulin molecule, CDR-3. This observation raises provocative questions concerning the nature and role of idiotypes in immune regulation. For example, if autologous antiidiotype recognized similar determinants, an apparently large number of antiidiotypes would be required to regulate this response, or auto-antiidiotype must recognize different determinants. Furthermore, it is unclear whether autologous antiidiotype can be made in vivo to determinants such as these which presumably occur on 'natural' antibody. Thus, while we have now begun to approach the molecular basis of idiotypy, additional questions have been raised as to the role of these markers in immune regulation.

Summary

A group of eight IgM hybridoma proteins induced with β(1,6)-D-galactan-containing antigens has been characterized in terms of primary amino acid sequence and idiotype expression. The H chain amino acid sequences reveal very strong homology in the V_H segment although several substitutions are seen that suggest the occurrence of somatic mutation in these IgM molecules. Significant sequence variation was observed in CDR-3, the region generated by the D segment, and the two recombination events, V_H-D and D-J_H. The number of amino acids in this region contributed by the D segment was found to vary from two to six, yet the overall length of CDR-3 was precisely maintained by the addition of amino acids on either side of D during the recombination processes. These additional amino acids are suggested to result from nucleotide addition by repair enzymes. Idiotypic analysis of these proteins, in conjunction with an assessment of the H chain sequences, has permitted an identification of the molecular basis of both cross-reacting and unique idiotypic determinants expressed by these molecules.

Received for publication 1 June 1983 and in revised form 26 July 1983.

References

1. Hood, L., E. Loh, J. Hubert, P. Barstad, B. Eaton, P. Early, J. Fuhrman, N. Johnson, M. Kronenberg, and J. Schilling. 1976. The structure and genetics of mouse immunoglobulins: an analysis of NZB myeloma proteins and sets of BALB/c myeloma proteins binding particular haptens. *Cold Spring Harber Symp. Quant. Biol.* 41:817.
2. Potter, M. 1977. Antigen-binding myeloma proteins of mice. *Adv. Immunol.* 25:141.
3. Seidman, J. G., A. Leder, M. Nau, B. Norman, and P. Leder. 1978. Antibody diversity. The structure of cloned immunoglobulin genes suggests a mechanism for generating new sequences. *Science (Wash. DC).* 202:11.
4. Cory, S., B. M. Tyler, and J. M. Adams. 1981. Sets of immunoglobulin V_κ genes homologous to ten cloned V_κ sequences: implications for the number of germline V_κ genes. *J. Mol. Appl. Genet.* 1:103.
5. Brack, C., M. Hirama, R. Lenhard-Schuller, and S. Tonegawa. 1978. A complete immunoglobulin gene is created by somatic recombination. *Cell.* 15:1.
6. Weigert, M., L. Gatmaitan, E. Loh, J. Schilling, and L. Hood. 1978. Rearrangement of genetic information may produce immunoglobulin diversity. *Nature (Lond.).* 276:785.
7. Max, E. E., J. G. Seidman, and P. Leder. 1979. Sequences of five recombination sites

encoded close to an immunoglobulin κ constant region gene. *Proc. Natl. Acad. Sci. USA.* 76:3450.

8. Sakano, H., K. Huppi, G. Heinrich, and S. Tonegawa. 1979. Sequences at the somatic recombination sites of immunoglobulin light chain genes. *Nature (Lond.).* 280:288.
9. Honjo, T., and T. Kataoka. 1978. Organization of immunoglobulin heavy chain genes and allelic deletion model. *Proc. Natl. Acad. Sci. USA.* 75:2140.
10. Rao, D. N., S. Rudikoff, H. Krutzsch, and M. Potter. 1979. Structural evidence for independent joining region gene in immunoglobulin heavy chains from anti-galactan myeloma proteins and its potential role in generating diversity in complementarity determining regions. *Proc. Natl. Acad. Sci. USA.* 76:2890.
11. Schilling, J., B. Clevinger, J. M. Davie, and L. Hood. 1980. Amino acid sequence of homogeneous antibodies to dextran and DNA rearrangements in heavy chain V-region gene segments. *Nature (Lond.).* 283:35.
12. Early, P., H. Huang, M. Davis, K. Calame, and L. Hood. 1980. An immunoglobulin heavy chain variable gene is generated from three segments of DNA: V_H, D, and J_H. *Cell.*19:981.
13. Sakano, H., R. Maki, Y. Kurosawa, W. Roeder, and S. Tonegawa. 1980. Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy chain genes. *Nature (Lond.).* 286:676.
14. Davis, M. M., K. Calame, P. W. Early, D. L. Livant, R. Joho, I. L. Weisman, and L. Hood. 1980. An immunoglobulin heavy chain gene is formed by at least two recombinational events. *Nature (Lond.).* 283:733.
15. Bernard, O., and N. M. Gough. 1980. Nucleotide sequence of immunoglobulin heavy chain joining segments between translocated V_H and μ constant region genes. *Proc. Natl. Acad. Sci. USA.* 77:3630.
16. Yaoita, Y., and T. Honjo. 1980. Deletion of immunoglobulin heavy chain genes from expressed allelic chromosome. *Nature (Lond.).* 286:850.
17. Sakano, H., Y. Kurosawa, M. Weiggert, and S. Tonegawa. 1981. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy chain genes. *Nature (Lond.).* 290:562.
18. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure and assembly of immunoglobulin heavy chain diversity DNA segments. *J. Exp. Med.* 155:201.
19. Rudikoff, S., D. N. Rao, C. P. J. Glaudemans, and M. Potter. 1980. κ chain joining segments and structural diversity of antibody combining sites. *Proc. Natl. Acad. Sci. USA.* 77:4270.
20. Weigert, M., R. Perry, D. Kelley, T. Hunkapiller, J. Schilling, and L. Hood. 1980. The joining of V and J gene segments creates antibody diversity. *Nature (Lond.).* 283:497.
21. Gough, N. M., and O. Bernard. 1981. Sequences of the joining region genes for immunoglobulin heavy chains and their role in generation of antibody diversity. *Proc. Natl. Acad. Sci. USA.* 78:509.
22. Weigert, M. G., I. M. Cesari, S. J. Yonkovich, and M. Cohn. 1970. Variability in the lambda light chain sequences of mouse antibody. *Nature (Lond.).* 288:1045.
23. Cesari, I. M., and M. Weigert. 1973. Mouse lambda chain sequences. *Proc. Natl. Acad. Sci. USA.* 70:2112.
24. Selsing, E., and U. Storb. 1981. Somatic mutation of immunoglobulin light chain variable region genes. *Cell.* 25:47.
25. Gearhart, P., N. D. Johnson, R. Douglas, and L. Hood. 1981. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature (Lond.).* 291:29.
26. Bothwell, A. L. M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D.

Baltimore. 1981. Heavy chain variable region contribution to the NPb family of antibodies: somatic mutation evident in a g2a variable region. *Cell.* 24:625.

27. Crews, S. J., J. Griffin, H. Huang, K. Calame, and L. Hood. 1981. A single V_H gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of antibody. *Cell.* 25:59.
28. Cook, W. D., S. Rudikoff, A. Giusti, and M. D. Scharff. 1982. Somatic mutation in a cultured mouse myeloma cell affects antigen binding. *Proc. Natl. Acad. Sci. USA.* 79:1240.
29. Rudikoff, S., A. M. Giusti, W. D. Cook, and M. D. Scharff. 1982. A single amino acid substitution altering antigen-binding specificity. *Proc. Natl. Acad. Sci. USA.* 79:1979.
30. Clarke, S. H., J. L. Claflin, and S. Rudikoff. 1982. Polymorphisms in immunoglobulin heavy chains suggesting gene conversion. *Proc. Natl. Acad. Sci. USA.* 79:3280.
31. Dildrop, R., M. Bruggemann, A. Radbruck, K. Rajewsky, and K. Beyreuther. 1982. Immunoglobulin V region variants in hybridoma cells. II. Recombination between V genes. *EMBO (Eur. Mol. Biol. Organ.) J.* 1:635.
32. Bentley, D. L., and T. H. Rabbits. 1983. Evolution of immunoglobulin V genes: evidence indicating that recently duplicated human V_κ sequences have diverged by gene conversion. *Cell.* 32:181.
33. Slater, R. J., S. M. Ward, and H. G. Kunkel. 1955. Immunological relationships among the myeloma proteins. *J. Exp. Med.* 101:85.
34. Oudin, J., and M. Michel. 1963. Une nouvelle forme d'allotypie des globulins du serum de lapin apparement liae a la function et a la specificite anticorps. *C. R. Hebd. Seances Acad. Sci. Ser. D Sci. Nat.* 257:805.
35. Jerne, N. 1974. Towards a network theory of the immune system. *Ann. Immunol. (Paris).* 125C:373.
36. Pawlita, M., E. Mushinski, R. J. Feldmann, and M. Potter. 1981. A monoclonal antibody that defines an idiotype with two subsites in galactan-binding myeloma proteins. *J. Exp. Med.* 154:1946.
37. Pawlita, M., M. Potter, and S. Rudikoff. 1982. κ-chain restriction in anti-galactan antibodies. *J. Immunol.* 129:615.
38. Jolley, M. E., S. Rudikoff, M. Potter, and C. P. J. Glaudemans. 1973. Spectral changes on binding of oligosaccharides to murine immunoglobulin A myeloma proteins. *Biochemistry.* 12:3039.
39. Zimmerman, C. L., E. Appella, and J. J. Pisano. 1977. Rapid analysis of amino acid phenylthiohydantoins by high performance liquid chromatography. *Anal. Biochem.* 77:560.
40. Jolley, M. E., C. P. J. Glaudemans, S. Rudikoff, and M. Potter. 1974. Structural requirements for the binding of derivatives of D-galactose to two homogeneous murine immunoglobulins. *Biochemistry.* 13:3179.
41. Mushinski, E. B., and M. Potter. 1977. Idiotypes on galactan-binding myeloma proteins and anti-galactan antibodies in mice. *J. Immunol.* 119:1888.
42. Ollo, R., C. Auffray, J. L. Sikorar, and F. Rougeon. 1981. Mouse heavy chain variable regions: nucleotide sequence of a germline V_H gene segment. *Nucleic Acids Res.* 9:4099.
43. Rodwell, J. D., and F. Karush. 1980. Restriction in IgM expression. I. The V_H regions of equine anti-lactose antibodies. *Mol. Immunol.* 17:1553.
44. Alt, F. W., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-J_H fusions. *Proc. Natl. Acad. Sci. USA.* 79:4118.
45. Clarke, S. H., J. L. Claflin, M. Potter, and S. Rudikoff. 1983. Polymorphisms in anti-

phosphocholine antibodies reflecting evolution of immunoglobulin families. *J. Exp. Med.* 157:98.

46. Rudikoff, S. 1983. Immunoglobulin structure-function correlates: antigen binding and idiotypes. *Contemp. Top. Mol. Immunol.* 9:169.
47. Mage, R., R. Lieberman, M. Potter, and W. D. Terry. 1973. Immunoglobulin allotypes. *In* The Antigens. M. Sela, editor. Academic Press, Inc., New York. 300–376.
48. Manjula, B. N., C. P. J. Glaudemans, and M. Potter. 1976. Subunit interactions in mouse myeloma proteins with anti-galactan activity. *Proc. Natl. Acad. Sci. USA.* 73:932.
49. Feldmann, R., M. Potter, and C. P. J. Glaudemans. 1981. A hypothetical space-filling model of the galactan-binding myeloma immunoglobulin J539. *Mol. Immunol.* 18:683.
50. Navia, M. A., D. M. Segal, E. A. Padlan, D. R. Davies, N. Rao, S. Rudikoff, and M. Potter. 1979. Crystal structure at 4.5 Å resolution of the galactan-binding mouse J539 immunoglobulin Fab. *Proc. Natl. Acad. Sci. USA.* 76:4071.
51. Morahan, G., C. Berek, and J. F. A. P. Miller. 1983. An idiotypic determinant formed by both immunoglobulin constant and variable regions. *Nature (Lond.).* 301:720.
52. Potter, M., E. B. Mushinski, S. Rudikoff, C. P. J. Glaudemans, E. A. Padlan, and D. R. Davies. 1979. *Ann. Immunol. (Paris).* 130C:263.
53. Vrana, M., S. Rudikoff, and M. Potter. 1979. The structural basis of a hapten-inhibitable κ-chain idiotype. *J. Immunol.* 122:1905.
54. Johnson, N., J. Slankard, L. Paul, and L. Hood. 1982. The complete V domain amino acid sequences of two myeloma inulin-binding proteins. *J. Immunol.* 128:302.
55. Clevinger, B., J. Schilling, L. Hood, and J. M. Davie. 1980. Structural correlates of cross-reactive and individual idiotypic determinants on murine antibodies to α(1,3) dextran. *J. Exp. Med.* 151:1059.
56. Clevinger, B., J. Thomas, J. M. Davie, J. Schilling, M. Bond, and L. Hood. 1981. Anti-dextran antibodies:sequences and idiotypes. *ICN-UCLA Symp. Mol. Cell. Biol.* 20:159.
57. Kabat, E. A., T. T. Wu, and H. Bilofsky. 1979. Sequence of immunoglobulin chains. National Institutes of Health. Publication 80-2008. p. 1.