THE PARTIAL AMINO ACID SEQUENCE OF THE HEAVY CHAIN VARIABLE REGION OF J558, A DEXTRAN-BINDING MYELOMA PROTEIN

Thesis by

Carol Oken

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

California Institute of Technology

Pasadena, California

1979

(Submitted January 9, 1979)

Introduction

The vertebrate immune system can recognize a huge number of antigens with the estimated 10^5 to 10^8 different antibody species it is capable of producing. Since 1900, immunologists have questioned how so much diversity is generated in the organism (1).

The antibody molecule is comprised of two types of polypeptide chains, termed light chains and heavy chains, which are joined by disulfide bridges. The polypeptide chains of different antibody species differ in their amino acid sequence (2) and the differences are largely contained in the N-terminal region of the chains called the variable region (3-7). The constant region is the remaining and relatively invariant (within an immunoglobulin class) portion of each chain. The variable and constant regions are coded for by different DNA sequences (8,9) which are rearranged during differentiation (10,11). There is evidence that the variable region itself is composed of two segments, the variable (V) segment and the joining (J) segment, each encoded by a separate DNA sequence (11,12). The specificity of an antibody molecule for its antigen or antigens is due to the complementarity of the antigen-binding site of the antibody and the antigenic determinant(s) (13). Each antigen-binding site is in a variable domain consisting of a heavy chain variable region (V_H) and a light chain variable region (V_L) (14-17). The primary structures of the variable regions determine the three-dimensional structure and therefore the antigen-binding properties of the variable domains. This was demonstrated in an experiment in which antigen-binding or F_{ab} fragments were first reduced to disconnect the disulfide linkages and denatured (18). Then, upon renaturation, the majority of the F_{ab} fragments folded back to their original conformation and thereby regained their original specificity.

Patterns of diversity have emerged from sequence analyses of many light chain variable (19,20) and heavy chain variable (5-7) regions. The heavy or light chain variable region sequences can be organized into subgroups having very similar sequences. Within these subgroups, the few amino acid differences are conservative and could result from single DNA base changes. By definition, there are relatively few sequence differences among members of the same variable region subgroup and many sequence differences in members of different variable region subgroups. It has been suggested that each subgroup is coded for by at least one germ line gene so the number of variable region subgroups is a minimal estimate of the number of germ line genes. Taking this as a premise, the origin of diversity within a subgroup remains unresolved.

Three major theories attempt to explain the origin of this diversity. According to the germ line theory, each of the variable regions is encoded by a different germ line gene. Therefore, the strict germ line theory predicts that all variable region diversity is inherited.

In the somatic theory, a small number of germ line variable genes (or, in its extreme form, one gene) can generate a large number of antibody variable regions by mutation and recombination during the somatic differentiation of antibody-producing cells. Since only the germ line genes are inheritable, the bulk of the diversity is separately and somewhat uniquely encoded in each individual.

The combinatorial joining model states that two or more DNA sequences, which are separated in the genome and brought together during the differentiation of an antibody-producing cell, code for a single variable region (11,12,21-24). An example of such a model proposes that two such DNA sequences, one coding for a V segment and the other coding for a J segment, are connected during

differentiation to form one variable region-encoding DNA sequence (11,12). In this version of the combinatorial joining model, the extent of diversity that an organism is capable of expressing is determined by its germ line V and J segment repertoire. The diversity observed is determined during somatic differentiation by the V and J segment combinations are expressed.

Antibody diversity has been studied by looking at subsets of the antibody repertoire of an organism. In particular, immunoglobulin subsets that bind known antigens are characterized. By sequencing sets of immunoglobulins that bind one antigen, structure-function relationships as well as antibody diversity patterns can be investigated.

Because of difficulties in the purification of antibodies elicited in an immune response, homogeneous immunoglobulins were first obtained from artificially induced tumors. In BALB/c mice, repeated intraperitoneal mineral oil injections result in neoplastic transformation of antibody-secreting cells producing myeloma tumors (25). Generally, for each mouse given this treatment a different myeloma tumor is induced and only one tumor is induced per individual. The immunoglobulin (myeloma protein) is secreted into the peritoneal cavity and the fluid that collects there, ascites fluid, is a good source of homogeneous immunoglobulin. The induced tumors can be stored (26) and transplanted into the peritoneal cavities of other mineral oil-primed mice (27). The specificities of an immunoglobulin produced by an induced myeloma tumor cannot be controlled by the experimenter but can be determined by screening with various antigens. In this way myeloma proteins with specificity for many different antigens (28-37, for example) including dextran B1355 (38) have been identified.

The use of myeloma immunoglobulins to investigate antibody diversity has been questioned since the distribution of the myeloma proteins that have been characterized is different from that of normal immunoglobulin. In the inbred BALB/c mouse strain, the percentage of the myeloma proteins that bind certain simple antigens such as dinitrophenol (DNP), phosphorylcholine (PC), and simple carbohydrates such as dextran (36) is larger than the estimated percentage of lymphocytes that bound the same antigens in an experiment using murine spleens (39). Also, a polypeptide chain having an unblocked N-terminus is thought to come from a different subgroup than one with a blocked N-terminus (40). Therefore, the observation that heavy chain N-termini are unblocked in about 80% of BALB/c myeloma proteins and are unblocked in only about 20% of serum immunoglobulins (41) is indicative of differing distributions of heavy chain variable regions in myeloma and serum immunoglobulins. In addition, in BALB/c mice, the amino acid sequences of some myeloma immunoglobulins have not been found in the normal immunoglobulin pool (42) and the sequenced normal immunoglobulins have members not found in the sequenced myeloma protein pool.

However, a myeloma tumor is thought to originate from a normal antibody-producing cell and secrete immunoglobulin that might normally be produced in the organism. The myeloma tumor inducing process may simply work on a set of antibody-producing cells which, because of their location, are restricted in specificity. This idea is supported by the fact that many antigens to which myeloma proteins bind are found normally in the murine intestine (43). There is evidence that myeloma proteins are the same as normal antibodies. The BALB/c immunoglobulins in the set of myeloma proteins that bind phosphorylcholine and dextran

B1355 are the same, at least using the criterion of idiotypy, as some of the antibodies induced by the same antigens in an immune response (44-46).

The question of whether study of the diversity in myeloma proteins is helpful in understanding the mechanisms of diversity in normal systems need not be dealt with at all now that lymphocyte hybridomas can be made. Using cell hybridization techniques, it has become possible to obtain hybridomas that secrete homogeneous antibody of predefined specificity (47,48). There are two routes to guanine nucleotide biosynthesis. The complete biosynthetic route is blocked by aminopterin. The scavenger route uses guanine or hypoxanthine as precursors and is catalyzed by hypoxanthine-guanine phosphoribosyl transferase (HGPRT). If 8-azoguanine is added to cells that have functional HGPRT, the drug is incorporated in the DNA and the cells die. Cells which are selected for 8-azoguanine resistance lack HGPRT, and will die on a medium containing hypoxanthine, aminopterin and thymidine (HAT medium) since they are forced to use the aminopterin-sensitive pathway which is blocked in this medium. However, if this type of cell is fused with a cell in which the HGPRT is functional, the hybrid cell will be able to survive on HAT medium.

To make hybridomas that are useful in antibody diversity studies, a non-immunoglobulin-secreting myeloma tumor cell line which lacks HGPRT is selected for by adding 8-azoguanine to the growth medium. These cells are unable to grow on HAT medium. Spleen cells from immunized mice also eventually die on HAT medium as do most somatic cell lines. Cells can be induced to fuse by adding polyethylene glycol (PEG). The resulting hybrid cells which have both functional HGPRT from the spleen lymphocyte and a rate of division approaching the myeloma cell survive on HAT medium. Clones of fused cells can then be assayed for their

binding specificity. The appropriate cell lines provide a good source of homogeneous antibody of the type secreted by the parental spleen cell (48). Amino acid sequence analyses of the immunoglobulins obtained in this way should give an excellent description of the kinds of diversity that exist in a set of variable regions which are of biological interest.

The Dextran System

The carbohydrate dextran B1355 elicits an immune response in mice that is relatively simple in several respects making it a good system for studying antibody diversity. Structural models for dextran binding immunoglobulins are provided by the three BALB/c myeloma proteins, M104 (μ,λ), J558 (α,λ), and U102 (α,λ) which have been identified as dextran binders by myeloma protein screening (49-51). Each of these dextran binding myeloma proteins is idiotypically distinct (though some idiotypic determinants are shared). Anti-idiotypic antisera, antisera made to recognize variable region antigenic determinants, have distinguished five types of antidextran antibody in the murine antidextran immune response. The four antisera were made to recognize determinants 1) unique to the M104 variable region, 2) unique to the J558 variable region, 3) unique to the U102 variable region, and 4) common to the variable regions of M104, J558, and U102. The five classes of dextran-binding antibody in the antidextran response that have been identified by the anti-idiotypic antisera are 1) IdI (M104), which has M104-like individual idiotypic determinants, 2) IdI (J558), 3) IdI (U102), 4) IdX, which has idiotypic determinants common to all three myeloma immunoglobulins, and 5) Id neg, which does not have idiotypic determinants that it shares with any of the three myeloma immunoglobulins (52). In these terms, the antidextran response

is quite a simple one, though some of the serologically identified antibody classes, especially IdX and Id neg, are likely to contain more than one antibody type.

The small number of spectrotypes observed in the isoelectric focusing patterns of murine 7S antidextran antibodies (the 19S antibodies cannot be focused on the gels used in the experiment) also indicates that the antidextran response is a limited one (46).

Isoelectric focusing demonstrates that the antidextran response in mice is not significantly changed in terms of the specificity and the idiotypic and size distribution of antibodies by an immunization procedure that considerably increases the quantity of antibodies produced. Mice are first immunized with dextran B1355 and hyperimmunized three weeks later. Then after one month, they are given three injections of Escherichia coli B which has antigenic determinants that are cross-reactive with dextran B1355. Due to the increase in antibody production, this immunization protocol permits analyses of the antidextran response in individual mice. Such analyses show that most of the major spectrotypes seen in isoelectric focusing patterns are shared among individual mice (46). The highly restricted isoelectric focusing patterns and the high degree of similarity among the patterns of individual mice indicate that germ line contributions to diversity may be important.

The dextran response appears to be controlled by a limited number of genes or gene clusters. Some inbred strains of mice, called high responders, respond to dextran B1355 immunization rapidly and appear to produce antibodies that are restricted in idiotype (53) and that carry similar if not identical lambda light chains. Light chains from dextran-binding antibodies in the serum of immunized mice cofocus on isoelectric focusing gels and focus with the light chains

of the dextran-binding myeloma proteins M104 and J558 (James Schilling, personal communication) which are known to be identical (54,50). Since only 3 to 5% of murine antibodies are lambda (55), as opposed to the other light chain class, kappa, the antidextran response is particularly specific with respect to light chain in high responders. Other inbred strains, called low responders, respond very slowly to dextran immunization and produce antibodies that lack both the idiotypes and the light chain found in the response of the high responders. The antibodies elicited in low responders have predominantly kappa light chains (53).

The antidextran response appears to be an autosomal dominant trait. The F_1 offspring of a cross between a high responder and a low responder all have the high responder phenotype (53). Recombination studies have shown the response is linked to heavy chain constant region markers (53).

The determination of the amino acid sequence of the heavy chain variable region of J558 was undertaken to gain information about idiotype, diversity, and the mechanism of antigen binding in the dextran system. The heavy chain of the myeloma protein M104 has been completely sequenced (56) and the M104 and J558 light chains are identical. The two immunoglobulins are idiotypically distinct (46,57) and they have identical light chains. The light chains can therefore be ignored as a source of variable domain heterogeneity and the focus can be placed on the heavy chain variable regions. Identifying the differences in the variable region of M104 and J558 (U102 is no longer available) should prove useful in idiotypic analyses of murine antidextran responses by making it possible to more precisely define the anti-idiotypic antisera. The nature of the differences in M104 and J558 sequences might also yield information about the mechanism of generating diversity in the dextran system.

The variable region sequence of the J558 heavy chain is also of interest as a model for dextran-binding immunoglobulins. Hypervariable regions are highly variable segments within antibody variable regions (21,58). Affinity labeling (59,60) and X-ray crystallography studies (61,62) have shown that these hypervariable regions form the antigen-binding crevice. One of the hypervariable regions in the M104 heavy chain contains a carbohydrate moiety attached to the amino acid backbone (56). M104 is the only sequenced immunoglobulin with known specificity having carbohydrate reported in a heavy chain hypervariable region (56). The heavy chain variable region sequence of J558 would help determine if the presence of carbohydrate in this hypervariable region is characteristic of dextran-binding immunoglobulins.

Experimental

Purification of J558. The BALB/c myeloma tumor J558 (Salk Institute) was carried subcutaneously in BALB/c mice and grown intraperitoneally in mineral oil-primed CDF1 mice (Cumberland View Farms). The ascites fluid was filtered twice through Miracloth (Chicopee Mills, Inc.), diluted with borate buffered saline (BBS) solution, pH 8.2, and the immunoglobulin was purified by affinity chromatography on a Sepharose-dextran column made using the procedure of S. March et al. (63) for cyanogen bromide activation and bovine serum albumin (BSA) coupling and the procedure of R. Hiramoto et al. (64) for dextran (given by M. E. Slodki, U.S. Department of Agriculture) coupling. Ascites fluid containing J558 immunoglobulin was passed through the column. The column was then washed with BBS, pH 8.2. The dextran-binding immunoglobulin was eluted from the column first with 0.1 M glycine-HCl buffer, pH 2.8 and then with 0.1 N HCl (64). The

homogeneity of the protein was monitored by electrophoresis on sodium dodecyl sulfate (SDS)-polyacrylamide gels (10% acrylamide) (65). The acid-eluted immunoglobulin was neutralized with 1 M NaOH and concentrated by ultrafiltration (Amicon).

Isolation of α Chain. The immunoglobulin was mildly reduced with dithiothreitol (10 mM) for 2 hours in EDTA and 0.2 M Tris under nitrogen. This was followed by alkylation with iodoacetamide (23 mM) for 1 hour on ice (66). The heavy and light chains were separated by gel filtration on an AcA 34 column (LKB) in 3 M guanidine and 0.2 M ammonium bicarbonate. The identity and homogeneity of the heavy chain was checked by polyacrylamide gel electrophoresis and the appropriate fractions were pooled, desalted by gel filtration on a G25 column (Pharmacia) in 0.2 M ammonium hydroxide, and lyophilized.

Isolation of Cyanogen Bromide Peptides. The α chains were dissolved in 70% formic acid and digested with cyanogen bromide (MC/B) at 4°C for 20 hours (67). The reaction mixture was diluted 1:10 with water and lyophilized. The cyanogen bromide fragments were separated by gel filtration on an AcA 54 column (LKB) in 3 M guanidine and 0.2 M ammonium bicarbonate.

The excluded volume fractions (CN1 peak) were fully reduced with dithiothreitol (10 mM) in guanidine (6 M) and Tris base in a 37°C water bath under nitrogen for 20 minutes. The fully reduced cyanogen bromide fragments were then alkylated with iodoacetamide (23 mM) for 1 hour on ice (68) and separated by gel filtration on an AcA 54 column.

The fractions in the CN1C peak (see Results section) were desalted by gel

filtration on a BioGel P2 column (BioRad) in 20% formic acid. The protein-containing fractions were diluted 1:2 with water and lyophilized.

The CN6 peak (sixth peak to elute) fractions from the separation of the heavy chain cyanogen bromide fragments were desalted on a P2 column in 20% formic acid, diluted 1:2 with water, and lyophilized.

Isolation of Tryptic Peptides. The polypeptides from the CN2 fractions were dissolved in 20% formic acid and dialyzed in 0.2 M ammonium bicarbonate. The fragments were digested by adding 1% (of volume of dissolved peptides) of a 10 mg/ml trypsin solution and incubating for 60 minutes at 37°C. Then 0.5% more of the trypsin solution was added and incubated at 37°C for 30 minutes. The 30-minute digestion was repeated twice and the solution was lyophilized. A two-dimensional separation of the tryptic peptides was performed using paper chromatography as the first dimension and high voltage electrophoresis as the second dimension (69). The solvents for the chromatography dimension were made by first mixing 800 ml water and 160 ml acetic acid in a 2 l separating funnel and mixing well. Then 540 ml n-butanol was added to the funnel and thoroughly mixed. The layers were allowed to separate for about 15 minutes and the aqueous (bottom) layer was used to equilibrate the chromatography tank for 1 hour prior to running the chromatography dimension. The organic (top) layer was used as the running solvent. The paper chromatography ran for 12 hours. The paper was then dried in an 80°C oven, dipped in the pyridine acetate buffer, pH 3.5, which is used for electrophoresis, and placed in a high voltage electrophorator (Gilson, Model D), which uses Varsol for cooling. The electrophoresis was allowed to run for 1 hour and the paper was dried in an 80°C oven. The tryptic peptides were

stained with ninhydrin and eluted from the stained spots, which were cut out of the two-dimensional map.

Composition and Sequence Determination. The tryptic peptides eluted from each spot on the two-dimensional map were hydrolyzed in a 1:1 solution of proprionic acid and hydrochloric acid for 20 minutes at 160°C. The amino acid compositions were determined using a Durrum D-500 amino acid analyzer.

Sequence analyses of the fractions from the CN6 peak and the CN1C peak were done on a modified Beckman sequenator (70,71) which performs automated Edman degradations. Polybrene R (Aldrich), a non-protein carrier which helps to hold proteins in the spinning cup, where the sequenator reactions take place, was used to load the peptide samples (72). The amino acid phenylthiohydantoin fractions from each degradation cycle of the sequenator were evaporated with a stream of nitrogen to remove the solvent. Each sample was then redissolved in acetonitrile and the phenylthiohydantoin derivatives were identified by high pressure liquid chromatography on a Dupont Zorbax ODS column (4.6 mm x 25 cm) using a Waters Associates chromatographic system (72).

Results

Purification of J558. The J558 whole immunoglobulin was purified from pooled ascites fluid by affinity chromatography. Absorbance profiles of the eluting of the immunoglobulin from the dextran-Sepharose column with glycine-HCl buffer and with HCl are shown in Figure 1. Polyacrylamide gel electrophoresis demonstrated that immunoglobulin was eluted with both these solutions, as seen in Figure 2. Slot A in Figure 2 contains a glycine-HCl-eluted sample and slot B contains an HCl-eluted sample. Since the samples undergo electrophoresis in a buffer containing

Figure 1. Absorbance Profile of the Dextran Affinity Purification of J558. After 1) allowing the ascites fluid to flow into the dextran-coupled Sepharose column and 2) washing the column with BBS (not shown), the dextran-specific immunoglobulin was eluted first with glycine-HCl buffer and then with HCl. Each arrow designates fraction at which indicated eluting was begun.

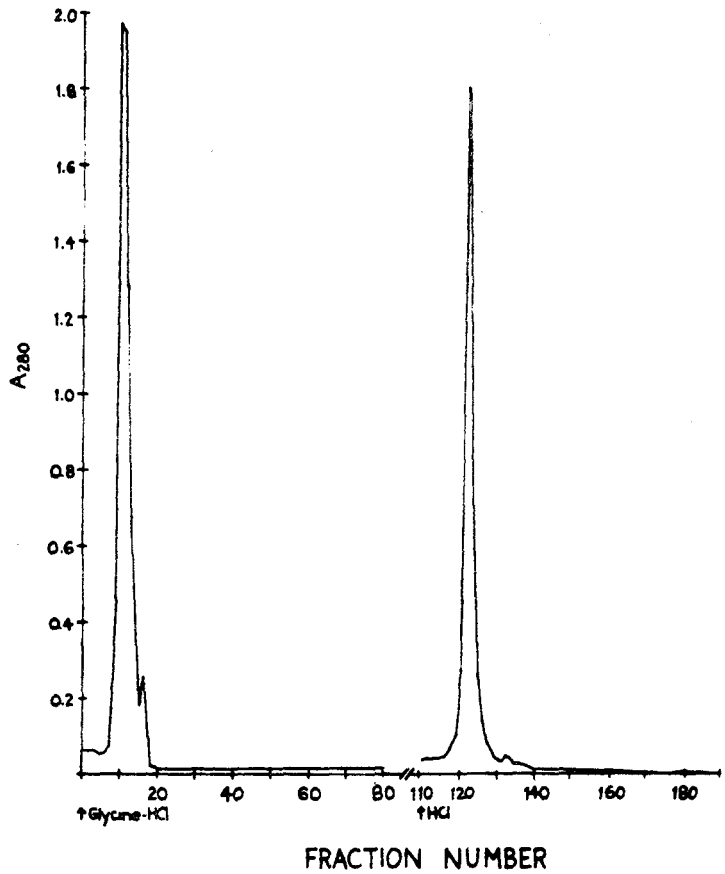
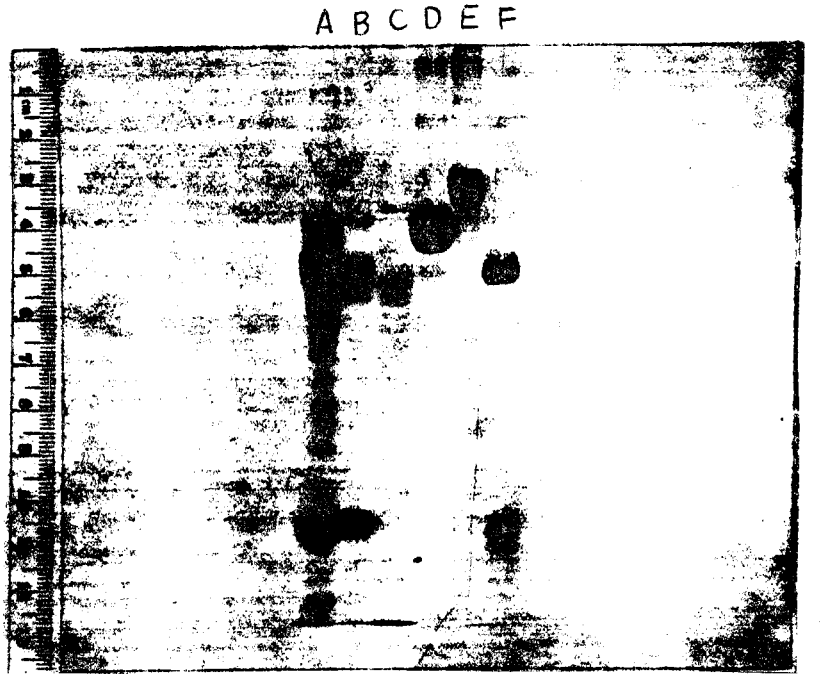


Figure 2. Polyacrylamide Gel Electrophoresis Showing Glycine-HCl-Eluted and HCl-Eluted Protein. (A) glycine-HCl-eluted protein, (B) HCl-eluted protein, (C) J558 (IgA standard), (D) bovine serum albumin (standard, MW 68,000), (E) M104 (IgM standard), (F) RAW116 (IgG standard).



2-mercaptoethanol, a reducing agent, interchain disulfide linkages are cleaved allowing the light and heavy chains to run separately. The molecular weight standards are IgA (J558) with heavy chain molecular weight of about 57,000, bovine serum albumin (BSA) with molecular weight of about 68,000, IgM (M104) with heavy chain molecular weight of about 65,000, and IgG (RAW116) with heavy chain molecular weight of about 50,000 in C, D, E, and F, respectively. The light chains for each immunoglobulin standard (IgA, IgM, and IgG) have a molecular weight of about 23,000. As the figure shows, the two major bands in both the glycine-HCl-eluted and the HCl-eluted samples run with the light chain (seen most clearly in slot F) and the J558 heavy chain (slot C). The same results were obtained upon gel electrophoresis of other samples taken throughout both the glycine-HCl-eluted and HCl-eluted peaks.

Isolation of α Chain. The whole immunoglobulin was first mildly reduced, a treatment which cleaves interchain disulfide bonds, and then alkylated to prevent the disulfide linkages from reforming. The heavy chains were separated from the light chains by gel filtration as shown in the absorbance profile in Figure 3. The two peaks were identified by polyacrylamide gel electrophoresis, as shown in Figure 4. Samples taken from five places throughout the first peak to elute are shown in slots 9 through 13. Each of these slots has one band which, by comparing with the standards, IgG (RAW116), IgM (M104), and IgA (J558), slots 1, 2, and 3, respectively, was identified as containing heavy chains. Each of the samples from the second peak, slots 4 through 8, has one band which runs with the light chain, as determined by comparison with the standards. The α heavy chain peak was pooled, desalted, and lyophilized.

Figure 3. Absorbance Profile of the Separation by Gel Filtration of J558 Heavy and Light Chains. Chain separation was performed on an AcA 34 column.

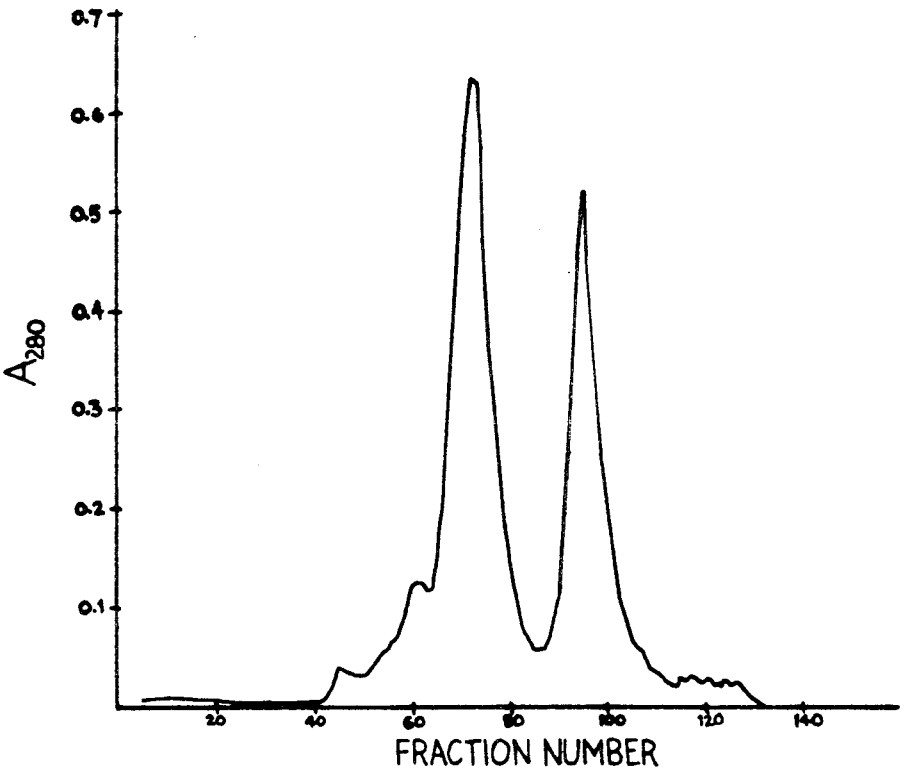
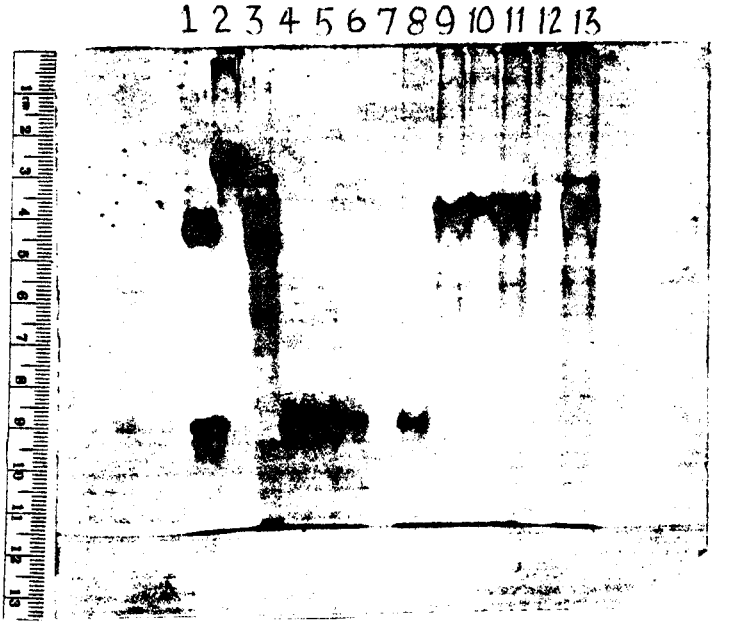


Figure 4. Polyacrylamide Gel Electrophoresis of Samples from the Two Peaks of the Chain Separation. (1) RAW116 (IgG standard), (2) M104 (IgM standard), (3) J558 (IgA standard), (4-8) samples from various points in the second peak to elute in chain separation, (9-13) samples from various points in first peak to elute in chain separation.



Isolation and Characterization of Cyanogen Bromide Peptides. The mildly reduced and alkylated heavy chains were digested with cyanogen bromide, which cleaves protein chains after methionine residues. The cyanogen bromide peptides produced by this treatment were separated by gel filtration as shown in the absorbance profile in Figure 5.

The sixth peak to elute, CN6, was pooled, desalted, lyophilized, and subjected to 5 cycles of degradation in the automatic sequenator. The yield of each amino acid phenylthiohydantoin at each degradation cycle is shown in Figure 6. Though these data are not conclusive since they do not point to a single sequence, the cyanogen bromide fragment containing the five residue sequence shown above the figure is evidently included in this sample. This sequence corresponds to residues 35 through 39 in the M104 heavy chain variable region. The contamination in the sample may be from other immunoglobulin cyanogen bromide peptides that elute with this peptide, from peptides of contaminating proteins present in the ascites fluid, or from randomly hydrolyzed peptides from either of these sources.

The first peak to elute, CN1, in the gel filtration separation of the cyanogen bromide peptides, Figure 5, was desalted, lyophilized, fully reduced to break intrachain disulfide bonds, and alkylated. The peptides generated were separated by gel filtration. The absorbance profile of this separation is shown in Figure 7. In another separation of IgA heavy chain cyanogen bromide fragments in which the same procedures were used, the peaks were labeled CN1A through CN1D in order of elution (73). In this separation, the CN1C peak (which contained a fragment which extended from residues 84 to 116) was the third to elute and eluted at about the same place as the peak eluting at about fraction 175 in Figure 7. Since the fragments from the fully reduced CN1 peak are likely to be the same

Figure 5. Absorbance Profile of the Separation by Gel Filtration of Cyanogen Bromide Peptides. An AcA 54 column was used to separate these fragments.

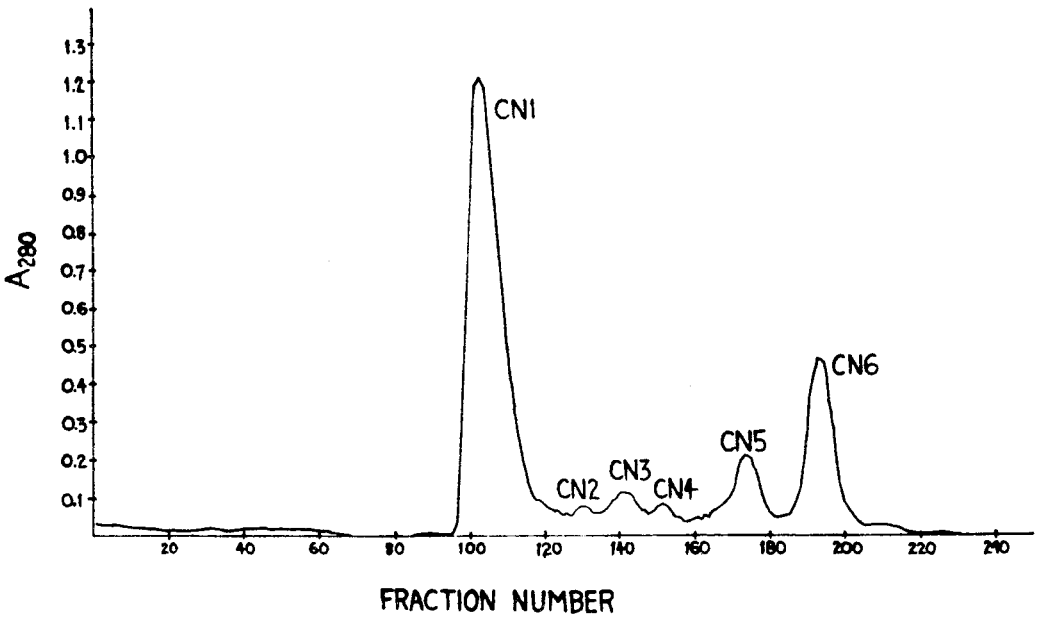


Figure 6. Yields of Amino Acid Phenylthiohydantoins from Five Sequenator Cycles of CN6 Cyanogen Bromide Peptide. The amino acids expected from this peptide at each sequenator step, based on the M104 sequence, are indicated in the graphs by large dots. The expected sequence is given above the figure. The expected sequence appears to be the predominant one in this preparation. However, cycle 1 appears to be especially heterogeneous with high yields of aspartic acid and glycine derivatives which do not fit with the expected sequence. Glycine yields are high throughout the run. The sequence above the figure has been recently confirmed by James Schilling (unpublished data).

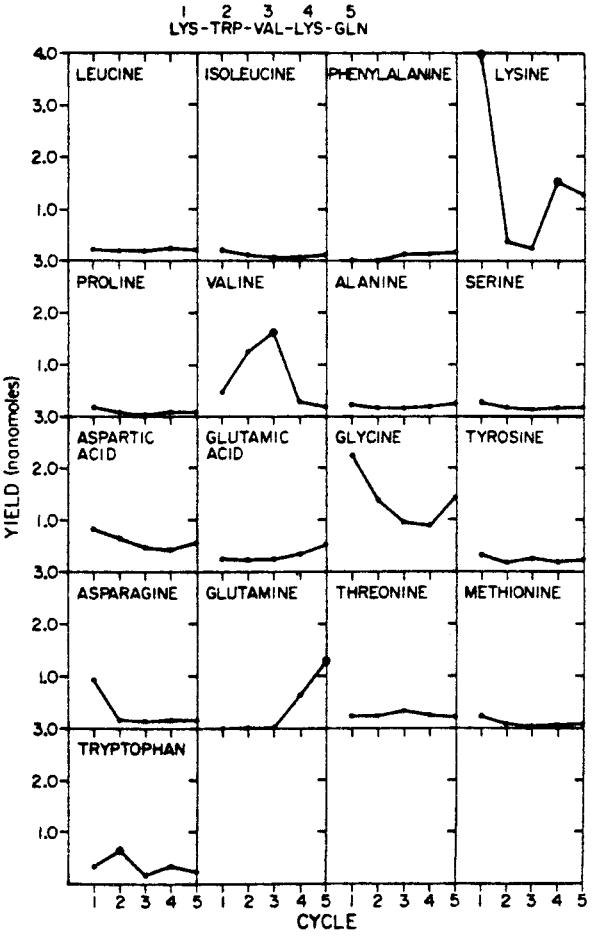
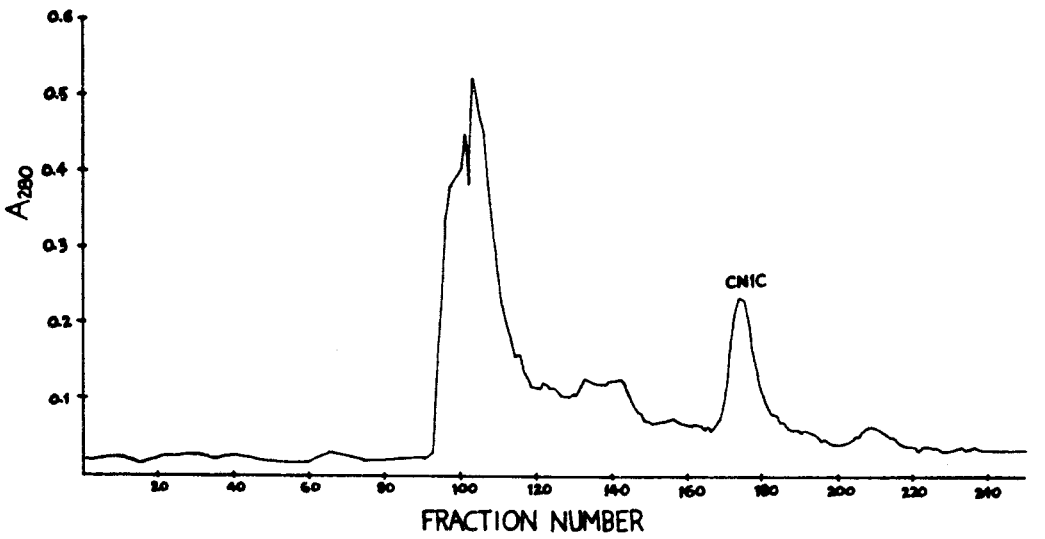


Figure 7. Absorbance Profile of the Separation by Gel Filtration of the Fully Reduced CN1 Peptides. An AcA 54 column was used to separate the fragments.



in both IgA molecules, the peak in the J558 separation was also labeled CN1C. This peak was pooled, desalted, and lyophilized. The yield of each amino acid phenylthiohydantoin from six cycles of degradation as well as the corresponding amino acid sequence are shown in Figure 8. Again, the sample is heterogeneous, but the fragment that was expected in this peak appears to be in the sample. The N-terminal six residues of this fragment are shown above the figure and corresponds to the M104 sequence from residue 82 to residue 87. The likely sources of contamination are the same as for the CN6 fragment, discussed above. In addition, the absorbance profile from fractions 90 to 160 was not very smooth and did not give the characteristic CN1A and CN1B peaks. The separation, therefore, may not have been good due to a poor column run.

Isolation and Characterization of Tryptic Peptides. A portion of the pooled CN6 peak was digested with trypsin, which cleaves after lysine and arginine, and the two-dimensional separation of the peptides is shown in Figure 9. Some of the ninhydrin spots were thought to contain variable region peptides because they migrated to approximately the same position as M104 variable region tryptic peptides. Since tryptophan fluoresces under UV, each UV positive spot indicated by broken lines on the map, contains tryptophan in its peptide. All spots were cut out and the peptides were eluted and hydrolyzed. The amino acid compositions of peptides believed to be from the variable region, based on the comparison with the M104 map discussed above, were determined. The amino acid analyses of each tryptic peptide from this fragment is shown in Table 1. These amino acid compositions correspond remarkably well to the tryptic peptide sequences for the corresponding M104 fragment (James Schilling, unpublished data).

Figure 8. Yields of Amino Acid Phenylthiohydantoins from Six Sequenator Cycles of CN1C Tryptic Peptide. The amino acids expected from this peptide at each step, based on the M104 sequence, are indicated in the graphs by large dots. The expected sequence appears to be included in the sample. High yields of valine in cycles 2 and 6, proline in cycles 1 and 2, and glycine in cycles 1, 2, and 5 do not fit with the expected sequence.

*The graph of serine gives the relative yield, based on absorbance at each cycle, of a serine product which could not be quantitated.

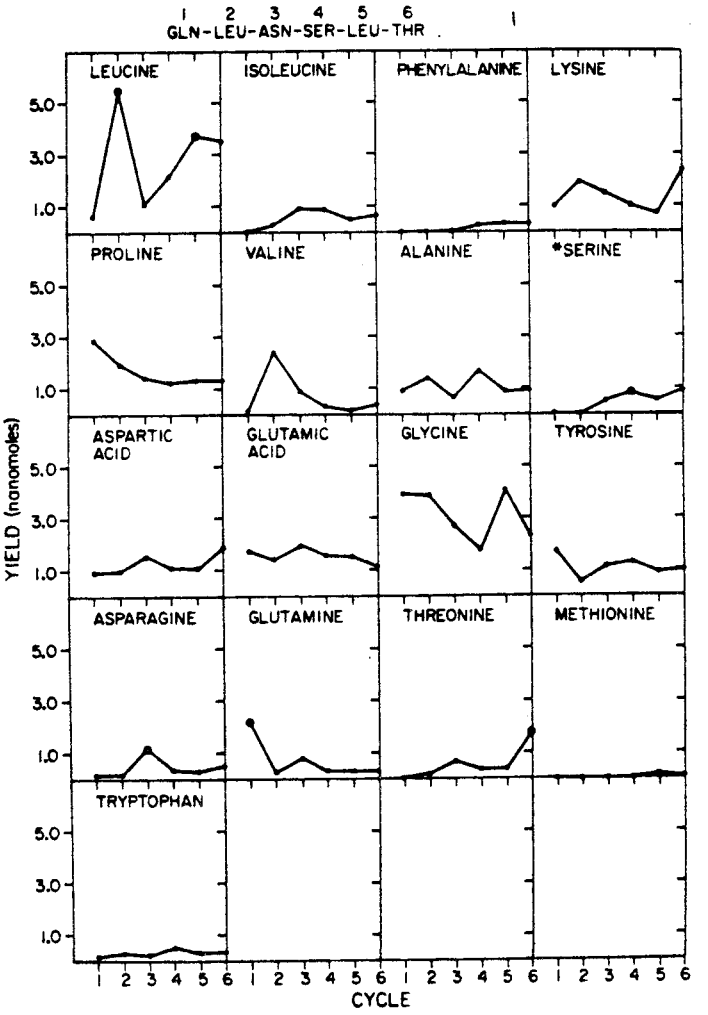


Figure 9. Two-Dimensional Map of Tryptic Peptides. Peptide spots are numbered in order of their appearance during staining. The peptides found to be fluorescent under UV are surrounded by broken lines.

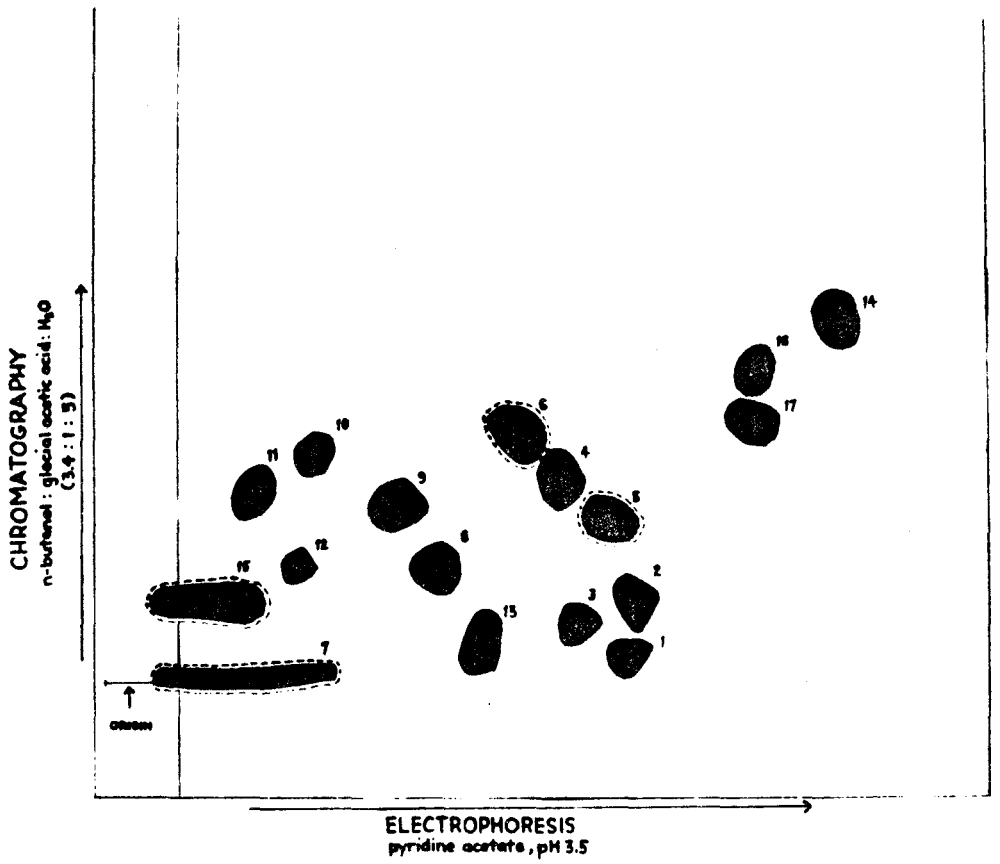


Table 1. Amino Acid Ratios of Tryptic Peptides. The tryptic peptides were generated from an aliquot of the pooled CN6 peak. The numbers in parentheses are the expected amino acid ratios based on the M104 sequence. The presence of homoserine indicates that the sample contains methionine.

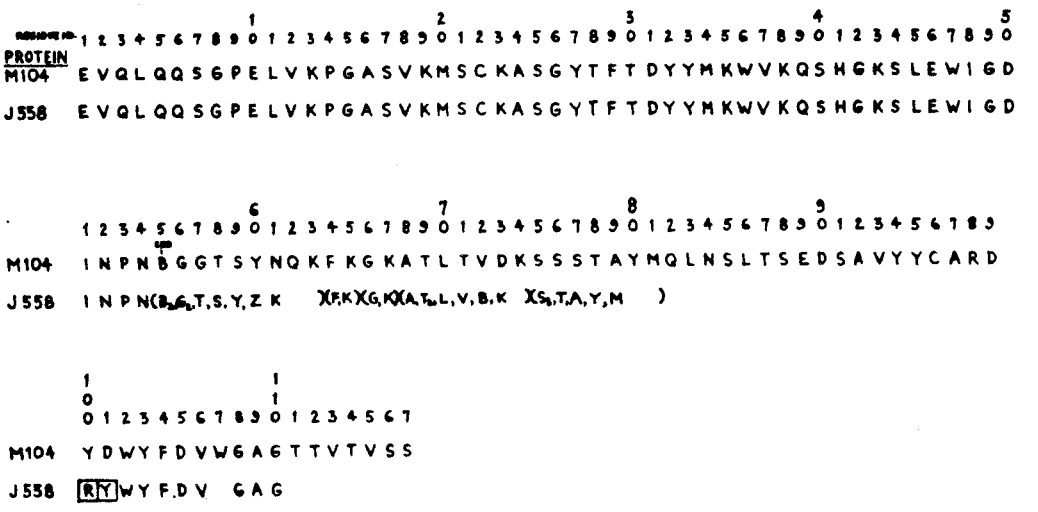
	PEPTIDE NO. (FROM PEPTIDE MAP, FIGURE 9)					
AMINO ACID(S)	1	2	4	9	12	15
ASP + ASN				1.15 (1)		4.65 (5)
THR				1.82 (2)	1.13 (1)	1.04 (1)
SER	0.98 (1)				2.65(3)	1.63 (2)
GLU + GLN	0.92 (1)					2.51 (2)
PRO						1.04 (1)
GLY	1.44(1)	0.72(1)				3.84 (3)
ALA				0.73(1)	1.30(1)	
VAL				0.87(1)		
ILE						2.13(2)
LEU				1.00(1)		1.17(1)
TYR					0.91(1)	0.98(1)
PHE			0.93(1)			
HIS	0.69(1)					
LYS	0.96(1)	1.28(1)	1.07(1)	0.71(1)		0.66(1)
HOMOSERINE					+	

Discussion

The dextran-binding myeloma protein J558 was successfully purified by affinity chromatography using dextran-coupled Sepharose columns. In order to determine the heavy chain variable region sequence for comparison with the myeloma protein M104, which also has specificity for dextran, variable region peptides were generated. The light and heavy chains of the immunoglobulin were separated by gel filtration after mild reduction and alkylation. The heavy chain was digested with cyanogen bromide and the peptides separated by gel filtration. One relevant peptide was obtained directly from this separation and the other was obtained after further reduction, alkylation, and separation by gel filtration.

The complete amino acid sequence of the M104 variable region (51) and the portions of the J558 variable region which have now been sequenced are shown in Figure 10. The N-terminal sequence was obtained by subjecting whole J558 heavy chain to automated stepwise degradation. In this way, the N-terminal 54 residues were identified by Hunkapiller and Hood (72) as was previously reported. The results of automated sequence analyses that have conclusively identified most of the CN1C fragment, residues 82 through 106 and residues 108 through 110 (James Schilling, unpublished data) are also shown in the figure. The variable region fragment which elutes in the CN6 peak appears to be identical by two criteria to the M104 cyanogen bromide fragment which extends from residue 35 to residue 81. The tryptic peptides from the J558 CN6 peak and those from the M104 cyanogen bromide fragment yield two-dimensional maps which are very similar. Furthermore, the amino acid compositions of the J558, shown in Figure 10, and M104 tryptic peptides from these fragments are identical.

Figure 10. The M104 Heavy Chain Variable Region Sequence and the Partial J558 Heavy Chain Variable Region Sequence. Residues 1-54 were determined by an N-terminal sequence analysis (72). Residues 100-106 and 108-110 were determined by James Schilling (unpublished data). The amino acid compositions thought to correspond to residues 55-81 are reported here.



As shown in Figure 10, there are two amino acid differences so far identified in the variable regions of M104 and J558. Both substitutions are located in the third hypervariable region which extends from residue 99 to residue 106. At residue 100, M104 has tyrosine and J558 has an arginine and at residue 101, M104 has an aspartic acid and J558 has a tyrosine. Both these changes are non-conservative and one requires a change of at least two DNA nucleotide bases of the three bases that code for the amino acid.

There appears to be a carbohydrate moiety in the J558 variable region located in the tryptic peptide that corresponds to the M104 variable region tryptic peptide which was found to contain carbohydrate. One spot on the two-dimensional map of the CN6 tryptic digest was found to have the same amino acid composition as the M104 carbohydrate-containing variable region tryptic peptide. This J558 spot was located at about the same place as the spot from the M104 peptide just discussed. Like the M104 peptide, this J558 peptide migrated only slightly in the chromatography direction and smeared in the electrophoretic direction, as is characteristic for carbohydrate-containing peptides. Also, in the sequenator run of the fragment starting at residue 35, the amount of phenylthiohydantoin recovered after the degradation cycle dropped significantly at residue 55 (James Schilling, unpublished observation). This has been observed at other residues at which carbohydrate is attached including residue 55 in M104. The presence of carbohydrate in this peptide can be verified by subjecting the peptide to a shorter hydrolysis and analyzing the hydrolyzate on the amino acid analyzer.

The anti-idiotypic antisera that have been made to J558 and M104 (46) indicate that each protein contains at least two idiotypic determinants, an IdI determinant, which is different in the two proteins, and an IdX determinant which

the two proteins share. Discussion of structural basis for and genetic origin of the idiotype differences in J558 and M104 must wait until the complete heavy chain variable region of J558 is determined.

It is now well established that a V segment DNA sequence and a J segment DNA sequence that are separated in the genome code for a whole variable region (9-11). The differences so far discovered in the heavy chain variable regions of M104, J558, as well as the dextran-binding hybridoma immunoglobulins currently being sequenced by James Schilling can be adequately accounted for by the germ line theory, the somatic mutation theory, or the combinatorial joining theory.

The M104 and J558 variable region sequences shown along with portions of several heavy chain variable region sequences of dextran-binding antibodies secreted from prepared hybridomas (James Schilling, unpublished data) suggest that the V segment ends at position 101. The variability within this segment appears to be restricted to residues 100 and 101, which are highly variable. The two amino acid sequence, starting with residue 101, is different in all seven variable regions in which these two residues have been identified due to many differences in both these positions. Based on this probable V segment-J segment boundary, three different J segments have been identified in the six dextran-binding heavy chains that have been sequenced in this region.

Each different V segment and J segment could be encoded by a different germ line sequence. Multiple V and J segment DNA sequences could then be attached in various combinations to generate the diversity observed. The fact that identical J segments are found attached to different V segments indicates that germ line contributions to diversity, at least in the J segment, are important.

Somatic mutation could generate all or part of the diversity observed. One V segment gene and one J segment gene may each undergo somatic mutation in a number of antibody-producing cells to give rise to the observed sequences. Alternately, there may be several V segment and/or J segment DNA sequences in the germ line and each of these may undergo mutation during differentiation. Somatic mutation may take place before or after the V and J segment DNA are joined to form a complete variable region-encoding DNA sequence. Much more variability could be generated if somatic mutation were an ongoing process and cells committed to one antibody species were capable of producing some progeny which displayed and could secrete antibody of altered specificity. This would provide a larger antibody pool for antigen selection. Since, as yet, identical V segments have not been found in the dextran-binding immunoglobulins characterized in this region, somatic mutation may play a role in V segment diversity.

The high degree of variability found in residues 100 and 101, on the V segment side of the hypothetical V segment-J segment junction, suggests that the mechanism by which V segment and J segment DNA sequences are joined generates the diversity. It is not yet known how the mechanism that leads to DNA rearrangement occurs, but if rearrangement takes place at or near the protein-encoding sequences, it may produce protein variability.

It is likely that each of these mechanisms contributes to the diversity observed in the dextran-binding myeloma proteins and antibodies. The relative contribution from each possible source remains to be determined. The V segment-J segment boundary can be confirmed by DNA studies like those that defined V segments and J segments in light chains (9-11). The specifics of the rearrangement of V and J segment DNA must await further experimentation at the DNA

level. Completion of the J558 J segment sequence and further sequence analyses of dextran-binding hybridoma antibodies will provide a more complete picture of the patterns of variability possible in dextran-binding immunoglobulins.

References

1. P. Ehrlich, Proc. Roy. Soc. Lond. 66, 424 (1900).
2. M. E. Koshland and F. M. Engleberger, Proc. Nat. Acad. Sci. U.S.A. 53, 1403 (1965).
3. N. Hilschmann and L. C. Craig, ibid. 53, 1403 (1965).
4. K. Titani, E. Whitley, Jr., L. Avogardo and F. W. Putnam, Science 149, 1090 (1965).
5. E. M. Press and N. M. Hogg, Nature 223, 807 (1969).
6. B. A. Cunningham, P. D. Gottlieb, M. N. Pflumn and G. M. Edelman, in Progress in Immunology, B. Amos, Ed. (Academic Press, New York, 1971), pp. 3-24.
7. B. A. Cunningham, M. N. Pflumn, U. Rutishauser and G. M. Edelman, Proc. Nat. Acad. Sci. U.S.A. 64, 997 (1969).
8. S. Tonegawa, N. Hozumi, G. Mathyssens and R. Schuller, Cold Spring Harb. Symp. Quant. Biol. 41, 877 (1976).
9. S. Tonegawa, A. M. Maxam, R. Tizard, O. Bernard and W. Gilbert, Proc. Nat. Acad. Sci. U.S.A. 75, 1485 (1978).
10. N. Hozumi and S. Tonegawa, ibid. 73, 3628 (1976).
11. C. Brack, M. Hirama, R. Lenhard-Schuller and S. Tonegawa, Cell 15, 1 (1978).
12. M. Weigert, L. Gatmaitan, E. Loh, J. Schilling and L. Hood, Nature 276, 785 (1978).
13. K. Landsteiner, The Specificity of Serological Reactions (Harvard Univ. Press, Cambridge, Mass., ed. 2, 1945).
14. F. Franek and R. S. Nezlin, Biokhimiya 28, 193 (1963).

15. G. M. Edelman, D. E. Olins, J. A. Gally and N. D. Zinder, Proc. Nat. Acad. Sci. U.S.A. 50, 753 (1963).

16. D. E. Olins and G. M. Edelman, J. Exp. Med. 119, 789 (1964).

17. S. J. Singer and R. E. Doolittle, Science 153, 13 (1966).

18. E. Haber, Proc. Nat. Acad. Sci. U.S.A. 52, 1099 (1964).

19. L. Hood, W. R. Gray, B. G. Sanders and W. J. Dreyer, Cold Spring Harb. Symp. Quant. Biol. 32, 133 (1967).

20. C. Milstein, Nature 216, 330 (1967).

21. T. Wu and E. Kabat, J. Exp. Med. 132, 211 (1970).

22. J. D. Capra and T. J. Kindt, Immunogenetics 1, 417 (1975).

23. E. A. Kabat, T. T. Wu, and H. Bilofsky, Proc. Nat. Acad. Sci. U.S.A. 75, 2429 (1978).

24. J. J. Cebra,T. H. Koo and A. Ray, Science 168, 263 (1974).

25. M. Potter and C. Boyce, Nature 193, 1086 (1962).

26. M. Potter, in Methods in Cancer Research, H. Busch, Ed. (Academic Press, New York, 1967), Vol. II, pp. 105-157.

27. M. Potter, J. G. Pumphrey and J. L. Walters, J. Nat. Cancer Inst. 49, 305 (1972).

28. M. Cohn, Cold Spring Harb. Symp. Quant. Biol. 32, 211 (1967).

29. H. N. Eisen, E. S. Simms and M. Potter, Biochemistry 7, 4126 (1968).

30. D. A. Schubert, A. Jobe and M. Cohn, Nature 220, 882 (1968).

31. B. J. Jaffe, H. N. Eisen, E. S. Simms and M. Potter, J. Immunol. 103, 872 (1969).

32. M. Potter and R. Lieberman, J. Exp. Med. 132, 737 (1970).

33. G. Vicari, A. Sher, M. Cohn and E. A. Kabat, Immunochemistry 7, 829 (1970).

34. M. Potter, Fed. Proc. 29, 85 (1970).

35. H. M. Grey, J. W. Hirst and M. Cohn, J. Exp. Med. 133, 289 (1971).

36. M. Potter, Ann. N.Y.Acad. Sci. 190, 306 (1971).

37. A. Sher, E. Lord and M. Cohn, J. Immunol. 107, 1226 (1971).

38. M. A. Leon, N. M. Young and K. R. McIntire, Biochemistry 9, 1023 (1970).

39. J. L. Press and N. Klinman, Eur. J. Immunol. 4, 155 (1974).

40. J. D. Capra and J. M. Kehoe, Adv. Immunol. 20, 1 (1975).

41. J. D. Capra, W. Wasserman and J. M. Kehoe, J. Exp. Med. 138, 410 (1973).

42. L. Hood, P. Barstad, E. Loh and C. Nottenburg, in The Immune System: Genes, Receptors, Signals, E. E. Sercarz et al., Eds. (Academic Press, New York, 1974), p. 119.

43. M. Potter, Physiol. Rev. 52, 631 (1972).

44. A. Sher and M. Cohn, Eur. J. Immunol. 2, 319 (1972).

45. D. Carson and M. Weigert, Proc. Nat. Acad. Sci. U.S.A. 70, 235 (1973).

46. D. Hansburg, D. E. Briles and J. M. Davie, J. Immunol. 117, 569 (1976).

47. G. Kohler and C. Milstein, Nature 256, 495 (1975).

48. R. C. Nowinski, M. E. Lostrum, M. R. Tam, M. R. Stone and W. N. Burnette, manuscript submitted.

49. K. R. McIntire, R. M. Asofsky, M. Potter and E. L. Kuff, Science 150, 361 (1965).

50. M. G. Weigert, I. M. Cesari, S. J. Yonkovich and M. Cohn, Nature 228, 1045 (1970).

51. J. Cisar, E. A. Kabat, J. Liao and M. Potter, J. Exp. Med. 139, 159 (1974).

52. D. Hansburg, D. E. Briles and J. M. Davie, J. Immunol. 119, 1406 (1977).

53. B. Blomberg, W. R. Geckeler and M. Weigert, Science 177, 178 (1972).

54. E. Appella, Proc. Nat. Acad. Sci. U.S.A. 68, 590 (1971).

55. K. R. McIntire and A. M. Rouse, Fed. Proc. 29, 704 (1970).

56. M. Kehry, J. Fuhrman, C. Sibley, J. Schilling and L. Hood, manuscript submitted.

57. M. Weigert, W. C. Rashke, D. Carson and M. Cohn, J. Exp. Med. 139, 137 (1974).

58. J. D. Capra, J. M. Kehoe, Proc. Nat. Acad. Sci. U.S.A. 71, 4032 (1974).

59. L. Wofsky, H. Metzger and S. J. Singer, Biochemistry 1, 1031 (1962).

60. D. Givol, Essays Biochem. 10, 73 (1974).

61. R. Poljak, L. M. Amzel, H. Avey, B. Chen, R. Phizackerley and F. Saul, Proc. Nat. Acad. Sci. U.S.A. 70, 3305 (1973).

62. L. M. Amzel, R. Poljak, F. Saul, J. Varga and F. Richards, ibid. 71, 1427 (1974).

63. S. C. March, I. Parikh and P. Cuatrecasas, Anal. Biochem. 60, 149 (1974).

64. R. Hiramoto, V. K. Ghanta, J. R. McGhee, R. Schrohenloher and N. M. Hamlen, Immunochemistry 9, 1251 (1972).

65. U. K. Laemmli, Nature 277, 680 (1970).

66. S. H. Bridges and J. R. Little, Biochemistry 10, 2525 (1971).

67. E. Gross, in Methods in Enzymology, S. P. Colowick and N. O. Kaplan, Eds. (Academic Press, New York, 1967), Vol. XI, pp. 238-255.

68. W. Konigsberg, in ibid., C. H. W. Hirs and S. N. Timashiff, Eds. (Academic Press, New York, 1972), Vol. XXV, pp. 185-188.

69. A. M. Katz, W. J. Dreyer and C. B. Anfinsen, J. Biol. Chem. 234, 2897 (1959).

70. B. Wittmann-Liebold, Hoppe-Seyler's Z. Physiol. Chem. 354, 1415 (1973).

71. B. Wittmann-Liebold, H. Graffunder and H. Kohls, Anal. Biochem. 75, 621 (1976).

72. M. W. Hunkapiller and L. E. Hood, Biochemistry 17, 2124 (1978).

73. S. Rudikoff and M. Potter, Biochemistry 13, 4033 (1974).