VARIABLE REGION SEQUENCES OF MURINE IgM ANTI-IgG MONOCLONAL AUTOANTIBODIES (RHEUMATOID FACTORS)

A Structural Explanation for the High Frequency of IgM Anti-IgG B Cells

By MARK J. SHLOMCHIK,* DAVID A. NEMAZEE,‡ VICKI L. SATO,§ JACQUES VAN SNICK,‖ DENNIS A. CARSON,¶ AND MARTIN G. WEIGERT*

*From the *Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111; the ‡Basel Institute for Immunology, CH4004 Basel, Switzerland; the §Biogen Research Corporation, Cambridge, Massachusetts 02142; the ‖Unit of Experimental Medicine, Universite Catholique de Louvain and International Institute of Cellular and Molecular Pathology, Brussels, Belgium; and the ¶Department of Basic and Clinical Research, Scripps Clinic and Research Foundation, La Jolla, California 92037*

IgM reactive with the Fc portion of self IgG can be elicited in the mouse by bacterial LPS (1–3), secondary protein immunization (4–6), or injection of immune complexes (7, 8). Secondary immunization also induces IgM anti-IgG in humans and rabbits (9, 10), and similar antibody has been found in rats (11). This autoantibody has been termed rheumatoid factor (RF)[1] because of its resemblance to IgM and IgG anti-IgG present in the serum of patients with rheumatoid arthritis (12). As an autoantibody response, the production of IgM anti-IgG is unusual in a number of respects: (*a*) it is not usually pathogenic or associated with a pathologic state; (*b*) it is a normal component of the secondary immune response, its production being stimulated by immune complexes formed during the response; (*c*) its production is regulated, peaking sharply at day 3 and falling to background in 1–2 wk; and (*d*) the percentage of B cells that can produce the antibody is strikingly high (1–8).

We have been intrigued by the high frequency of B cells that can make IgM anti-IgG. As many as 3–15% of hybridomas induced by the polyclonal activator LPS produce this type of antibody, most of which are specific for IgG1 (reference 3; D. Carson, unpublished results; J. Van Snick, unpublished results). The same frequency has been observed among hybrids elicited specifically via secondary immunization with protein antigens (5). The fraction of anti-IgG clones among LPS hybridomas is much higher than the fraction that binds to a variety of other antigens (13–18).

This work was supported by National Institutes of Health grants GM-20964-12 to M. Weigert and AM-25443 to D. Carson, and by an appropriation from the Commonwealth of Pennsylvania. M. Shlomchik is a trainee of the Medical Scientist Training Program at the University of Pennsylvania (NIH grant 5-T32-GM-07170).

[1] *Abbreviations used in this paper:* CDR, complementarity determining region; cOVA, chicken OVA; dNTP, deoxyribonucleic triphosphate; FR, framework region; RF, rheumatoid factor; sssDNA, sonicated salmon sperm DNA.

There are three plausible models to explain the high frequency of anti-IgG B cells. The "multiple epitope" model posits that there is an unusually large number of antigenic epitopes on IgG Fc, each with an ordinary frequency of B cells that can recognize it. A prediction is that sequences of anti-IgG antibodies will appear random since these antibodies are directed against a mixture of a large number of different epitopes. The "constant stimulation" model hypothesizes that constant exposure to autoantigen leads to vast clonal expansion of an ordinary number of precursors. The sequences of B cells that are derived from greatly expanded clones are expected to harbor large numbers of somatic mutations, since this process accompanies clonal expansion. The premise of the "combinatorial" model is that the presence on an antibody of a single V region structure (e.g., a particular V_H, V_κ, J_H, etc.) is sufficient to confer the ability to bind IgG Fc. Since such single structures would be expressed in a high percentage of B cells, this model suggests a possible structural explanation for a high precursor frequency of anti-IgG B cells.

Since these models make different predictions about the nature of V regions of IgM anti-IgG1 antibodies, we sought to distinguish between them by determining the sequences of such V regions expressed in monoclonal IgM anti-IgG1 antibodies. The data obtained support the combinatorial model and contradict the constant stimulation and multiple epitope models. In addition, the sequences suggest that the IgM anti-IgG1 autoantibodies may be binding IgG1 through an unusual combining site.

Materials and Methods

Hybridomas. Each hybridoma is derived from a separate fusion, except RF49 and RF51. The "JV" hybridomas were produced by fusion with SP2/0 and tested essentially as described (3, 5). JV6 is derived from a secondary immunization with DNP–human apolactoferrin. The "RF" hybridomas were produced from mice given 50 μg of LPS 3 d before fusion. Splenocytes from these mice were hybridized either with Ag8 (RF34) or SP2/0 (RF49 and RF51). Hybridomas were screened for binding of IgG of all mouse isotypes in an ELISA which detected bound IgM. Hybrids were subcloned and retested before use. VS1 and VS2 were produced from mice immunized with KLH. VS3 and VS4 were made from splenocytes of LPS-immunized mice. SP2/0 is the fusion partner in "VS" hybridomas.

RNA Preparation. Total cellular RNA was prepared by the guanidinium isothiocyanate method, essentially as described by Chirgwin et al. (21). Poly(A)$^+$-RNA was isolated on oligo(dT)-cellulose columns by standard methods (22).

Oligonucleotide Primers. κ and μ oligonucleotide primers were synthesized using phosphotriester chemistry either manually or by an oligonucleotide synthesizer (Applied Biosystems, Inc., Foster City, CA). Oligonucleotides were purified by electrophoresis through 40-cm 20% polyacrylamide/7 M urea preparative gels, followed by passive elution and Sephadex-DEAE chromatography. The sequences are (κ) 5′-TGGATGGTGGGAAGATG-3′, and (μ) 5′-GCAGGAGACGAGGGGGA-3′.

cDNA Synthesis and Purification. 50 ng of oligonucleotide primer was labeled with ^{32}P at the 5′ end with 10 U T4 polynucleotide kinase (Pharmacia Fine Chemicals, Piscataway, NJ) for 1 h at 37°C in the following reaction mix: 50 mM Tris-Cl (pH 7.6), 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, and 10 mM $MgCl_2$. The reactions were boiled for 2 min to inactivate the kinations enzyme. Kinations were then added to 30–60 μg of poly(A)$^+$-RNA and the mixture was heated to 90°C for 2 min. The mixtures were allowed to slow-cool for 20 min. The RNA-kination mixture was then adjusted to 100 mM Tris-Cl (pH 8.3), 140 mM KCl, 10 mM $MgCl_2$, 0.5 mM deoxyribonucleic triphosphates (dNTPs) (PL Biochemicals, Milwaukee, WI), 10 mM DTT, 120 U RNasin (Promega Biotec,

Madison, WI), 40 U reverse transcriptase (Life Sciences, St. Petersburg, FL), and cDNA synthesis was allowed to take place for 2 h at 42°C. The reaction was stopped by adjustment to 150 μg/ml boiled RNAase (Sigma Chemical Co., St. Louis, MO), 15 mm EDTA, and further incubation at 42°C for 30 min. The reaction was then extracted once with an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol and then back-extracted with an equal volume of Tris/EDTA (10:1, pH 8.0). Aqueous phases were pooled and brought to 0.3 M sodium acetate with 4 M sodium acetate (pH 6.0) and precipitated with ethanol. Dried pellets were resuspended in dH20 and formamide dyes and run on a 5% polyacrylamide 7 M urea gel for 2–3 h at 30 vol/cm. The gel was exposed to film for 5–10 min and the exposure was used to locate and excise relevant full-length cDNA transcripts, which usually contained the vast majority of incorporated radioactivity. Gel slices were chopped and incubated, shaking at 37°C overnight in a solution of 0.5 M ammonium acetate with 1 mM EDTA. The aqueous phases were drawn off and passed over a column of siliconized glass wool. 6.4 μg of sonicated salmon sperm DNA (sssDNA) was added and the solutions were precipitated with 2.5 vol ethanol. Pellets were redissolved in 0.3 M sodium acetate (pH 7.0) and reprecipitated with 2.5 vol ethanol. Pellets were then washed twice with 70% ethanol, once with 100% ethanol, and then dried. Finally, the pellets were redissolved in an appropriate volume of dH20 before aliquotting for sequencing.

Modified Chemical Degradative Sequencing. Reactions were modified from Rubin and Schmid (23) and Bencini et al. (24) as follows: 10 μl of labeled cDNA solution was aliquotted to each of four microfuge tubes, labeled A > C, G + A, C, T > G. 1 μl of sssDNA was added to the C and T > C tubes and the following reactions were carried out: A > C, add 2 μl of 3.75 M NaOH, incubate at 90°C for 9 min, cool on ice 3 min, add 150 μl 1 M piperidine; G + A, add 2 μl 0.5 M sodium formate (pH 2.0), incubate at 37°C for 15 min, add 150 μl 1 M piperidine, and keep on ice; C, heat 2 min at 90°C, cool 2 min on ice, add 20 μl of 4 M NH_2OH (pH 6.0 with diethylamine), incubate 23°C for 8 min, stop by adding 400 μl of cold 0.3 M sodium acetate, pH 7.0, and 20 μg/ml tRNA and 1 ml ethanol, freeze on dry ice for 5 min, and spin in a high speed microcentrifuge for 10 min. Resuspend in 300 μl 0.3 M sodium acetate, pH 7.0, and add 900 μl ethanol, freeze, and spin again. Wash once with 70% ethanol, once with 100% ethanol, and dry. Resuspend in 150 μl 1 M piperidine; T > G, heat 2 min at 90°C, cool 2 min on ice, add 20 μl of 0.04 mg/ml $KMnO_4$ and incubate at 23°C for 8 min, stop by adding 10 μl allyl alcohol, dehydrate by adding 1 ml 1-butanol, vortexing, and spinning for 2 min, resuspend pellet in 150 μl dH20 and add 1 ml butanol, vortex, and spin, dry, and resuspend in 150 μl 1 M piperidine. Incubate all reactions at 90°C for 30 min, cool on ice for 3–4 min, add 150 μl 70% ethanol to the A tube, and fill microfuge tubes almost to volume with 1-butanol, vortex, and spin for 2 min. Remove supernatants and add 150 μl 1% SDS to pellets, add 1 ml 1-butanol, vortex, spin 2 min, remove supernatants, add 1 ml ethanol, spin 3 min, remove supernatants, and dry. Resuspend in dH20 and formamide dyes and run on 80-cm 5% polyacrylamide linear gels and 80-cm 6% buffer-gradient polyacrylamide gels (25). In contrast to direct dideoxy-sequencing of mRNA (26), partial chemical degradation of a full-length cDNA nearly always allows the unambiguous determination of all bases up to ~400 nucleotides from a single primer hybridization site.

Results

IgM Anti-IgG1 Light Chain Sequences

The light chain nucleotide sequences from nine IgM hybridomas that bind IgG1 have been determined; they are summarized in Table I and presented as amino acid translations in Fig. 1. There are 10 light chain sequences from 9 hybridomas because RF49 synthesizes two very similar light chain mRNAs, neither of which has any defects in the sequenced region, including the leader sequences (data not shown), that would make them untranslatable. RF49 also produces two V_κ proteins, both of which associate with heavy chain in binding

TABLE I

BALB/c IgM Anti-IgG (RF) Hybridoma Characteristics Summary

Hybridoma*	Specificity	Induction	V_κ group	J_κ	V_H group	J_H	D‡	CDR3 length
VS1	IgG1	Secondary response to protein Ag	1	1	J606	3	Q52	8
VS2	IgG1	Secondary response to protein Ag	19	1	7183	1	FL16.1	10
VS3	IgG1	LPS	8	2	J558	4	?	8
VS4	IgG1	LPS	8	1	J558	3	FL16.1 or .2	7
RF49s	IgG3 > IgG1	LPS	19 (A chain)§ 19 (B chain)	2 1	Not known	—	—	—
RF51s	IgG1	LPS	19 (MPC11 gene)	2	J558	2	SP2.3, .4, .6	8
RF34	IgG1	LPS	19 (MOPC21 gene)	2	J558	4	any SP2	8
JV2	IgG1	LPS + dextran	8	1	J558	2	?	7
JV6	IgG1	Secondary response to protein Ag	1	1	J558	4	FL16.1 or .2	9

V region group assignments were made on the basis of >80% nucleic acid homology to prototype sequences.

* The VS series was produced by Sato and Nemazee, the RF series by Carson, and the JV series by Van Snick.

‡ All germline D_H genes that the sequenced D_H could have been derived from are given.

§ RF49 makes two light chains, both of which are V_κ 19 type and both of which associate with the heavy chain to bind IgG_1 (see text). The A chain is the gene used in MPC11 and is identical in sequence to RF51.

IgG. The number of restriction fragments detected with C_κ and J_κ probes in Bam HI– and Hind III–digested RF49 DNA indicates that RF49 is probably not the result of two spleen cells fusing with one SP2/0 cell (data not shown).

The light chains fall into only 3 of 24 known "families," according to the classification scheme proposed by Potter et al. (27). This system groups closely related sequences into families based on the relatedness of *N*-terminal amino acid sequences. The families identified in the IgM anti-IgG1 sequenced here are: $V_\kappa 1$ (two examples), $V_\kappa 8$ (three examples), and $V_\kappa 19$ (five examples).

An apparent bias in V_κ gene usage among RFs prompted us to calculate whether the light chain usage in RFs deviated significantly from random selection from the pool of expressed V_κ genes. For this calculation, it is necessary to know the fraction of expressed V_κ which each V_κ group comprises in unselected adult BALB/c B cells. This information has not been determined directly; however, a source of suitable data comes from two types of studies.

The analysis of Ig variable regions produced by independently induced plasmacytomas and LPS-derived hybridomas is one approach to sampling of the normally expressed pool of V regions. Surveys of this type have provided detailed information on the relative expression of many V_κ families (27–29, see below). Although some biases may exist in the genesis of plasmacytomas, which could affect the V repertoire being expressed in such tumors, the V region representation generally parallels that of serum Ig (30).

The germline gene composition of each V_κ family is a second source (31, 32). The assumption in applying this data is that, in terms of gene families, the selection of germline genes for expression is essentially unbiased. Several experimental approaches have found that, in adults, V gene usage is approximately commensurate with representation in the germline for most V genes (18, 28, 33–37).

To provide two different estimates of whether V_κ genes used in RFs are a random sampling of all V_κ genes, we have made calculations based on data from both the germline pool, as measured by probing Southern blots of germline

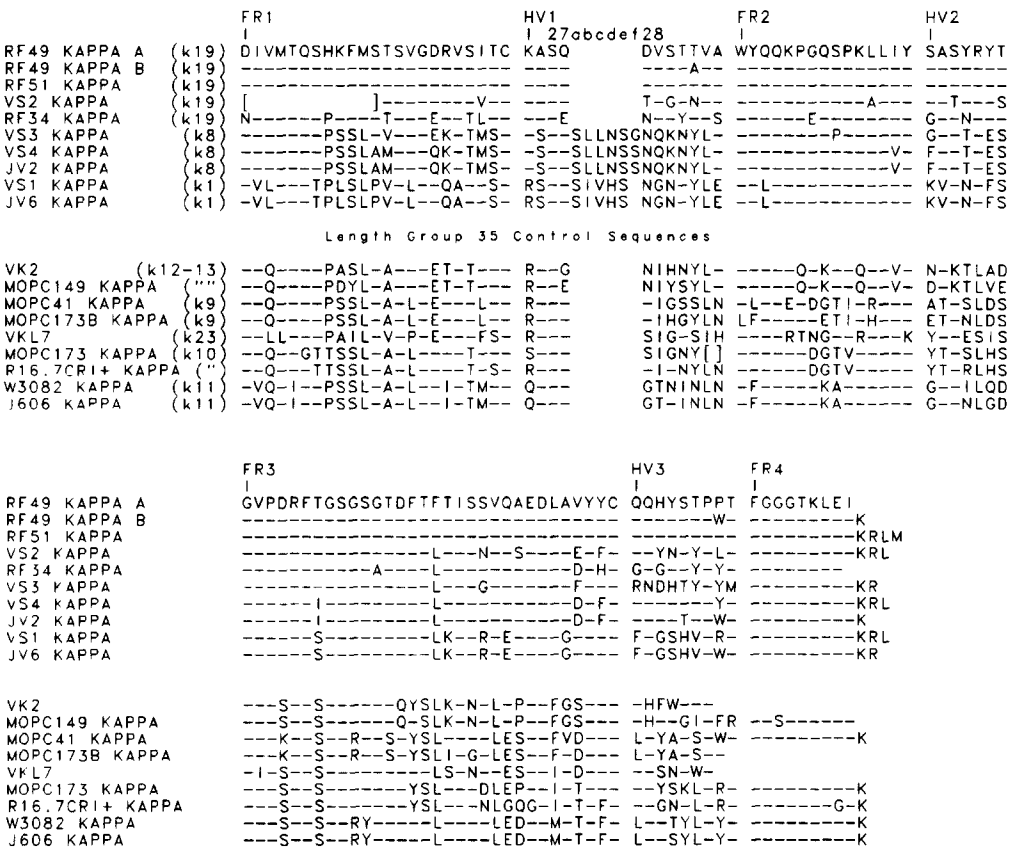

FIGURE 1. IgM anti-IgG1 (RF) V_κ amino acid sequences and control V_κ sequences. The RF amino acid sequences are inferred from the complete mRNA nucleotide sequences. Dashes indicate identity with the topmost sequence, which is given explicitly. Blank spaces are used to align sequences of different lengths for maximum homology. Framework (*FR*) and hypervariable (*HV*) regions, as defined by Kabat et al. (47), are indicated. The V_κ group to which each sequence belongs, as defined by Potter et al. (25), is given in parenthesis after the name. The *N*-terminal 12 residues of VS2 have not been determined and are indicated by empty brackets, as are two missing residues from MOPC173B. RF49 makes two productive V_κ chains (see text); they are given as "RF49 Kappa A" and "RF49 Kappa B."

DNA with V_κ probes, and the expressed pool, as measured by sequencing and electrophoretic analysis of proteins made by plasmacytomas. Estimates for the total pool size of germline (31) and expressed (27, 28, 31, 35) genes are similar and are in the range of 100 to 300. It should be noted that these values are likely to be underestimates (see discussion in references 27, 28, and 35).

Table II shows the results of these calculations. The different estimates of the fraction of the germline and of the expressed V_κ genes which each group comprises and the corresponding p values are given. The sizes of each of the V_κ families are derived as follows: $V_\kappa 19$ probes hybridize to seven bands in Southern blots of BALB/c liver DNA (31, 38). Assuming that the average band contains one gene (i.e., that each band represents one unique hybridizing restriction fragment and that each restriction fragment contains one gene) (31, 39, 40), $V_\kappa 19$ represents at most (since some bands may represent psuedogenes [41, 42])

TABLE II

The Significance of Restricted V_κ Usage in the RF by V_κ Group Based on Different Estimates of Group Representation in the V_κ Pool

V_κ group	Number of times observed	Estimated fraction of V_κ gene pool	Basis of estimate	p value
1	2	0.02	2 per 100 germline genes	0.014
		0.06	Percent of BALB/c myelomas expressing V_κ1C only	0.10
		0.12	Percent of BALB/c myelomas expressing V_κ1C and V_κ1A	0.30
		0.04	Maximum fraction of pool for $p < 0.05$ with this data	~0.05
8	3	0.08	8 per 100 germline genes	0.03
		0.02	Percent of BALB/c myelomas expressing V_κ8	0.002
		0.10	Maximum fraction of pool for $p < 0.05$ with this data	~0.05
19	4	0.07	7 per 100 germline genes	0.002
		0.02	Percent of BALB/c myelomas expressing V_κ19	0.00002
		0.17	Maximum fraction of pool for $p < 0.05$ with this data	~0.05

Different estimates of the representation of the V_κ groups in the pool of available V_κ genes and the respective p values based on those estimates are given. The null hypothesis under test is that, for each individual V_κ family, the representation in the sample is consistent with random selection from the general population. Therefore, p values are the binomial probabilities of selecting at random a sample containing the observed number of hybridomas expressing the light chain from the given V_κ group from the general population. The composition of the general population is not exactly known, and so, as discussed in the text, is estimated from data derived from several experimental approaches. Several estimates are used to show the effect that error in the estimate would have on our conclusions. The estimates described in the text are referred to in the column "Basis of estimate." Briefly, the number of germline genes in each family is based on Southern blotting data. The total number of V_κ genes is conservatively assumed to be 100, based on a minimum value from several estimates. The percentage of expression in BALB/c myelomas is taken from reference 28. See text for details and further references.

7 out of 100 of the germline pool. With a V_κ8 probe, six to eight bands were detected by Southern blotting (reference 31 and L. D'Hoostelaere, personal communication). For V_κ1, two bands are detected in a Southern blot (32 and L. D'Hoostelaere, personal communication). BALB/c plasmacytoma libraries contained 2% V_κ19-, 2% V_κ8-, and 6–12% V_κ1-expressing tumors (28). The V_κ1 gene family seems to be one of only 3 out of 65 groups that was overrepresented. In the NZB tumors, the V_κ1 family, which is encoded in the NZB genome by a different set of highly homologous genes (32), was not overrepresented (28).

Using the germline gene estimate for the fraction of the pool that each group occupies, the p values indicating nonrandom expression of the groups are: V_κ1, 0.014; V_κ8, 0.03; and V_κ19, 0.002. Thus, a limited set of V_κ genes, roughly representing 5–17% of germline V_κs and 10–16% of expressed V_κs, is found in IgM anti-IgG hybridomas. Considering the selection of V_κs as a whole, if each of the 24 V_κ families were equally likely to be selected in RFs, then the chance that three or fewer different families would be seen in nine trials is 1.4×10^{-5}.

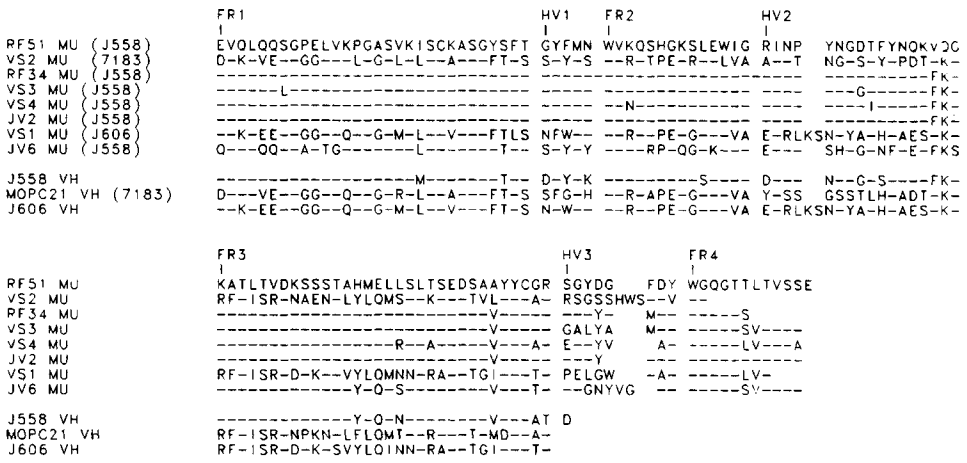

FIGURE 2. IgM anti-IgG1 (RF) μ chain sequences. Sequences are derived and presented as in Fig. 1. The V_H sequences of J558, MOPC21, and J606 (taken from reference 47) are presented for comparison. The variable-length CDR3 sequences have been aligned for maximum homology.

V_H Gene Sequences

The heavy chain sequences for eight IgM anti-IgG were determined and are presented as amino acid translations in Fig. 2 and summarized in Table I. The RF49 heavy chain was not sequenced because the cell line lost production of heavy chain mRNA during propagation before preparation of RNA. The same type of analysis applied to the light chains to test if they constitute a randomly selected set can also be applied to the heavy chains. There are eight described V_H gene families constituting 100–200 genes (39, 43). Heavy chain gene families have been characterized in the germline as sets of nonoverlapping bands, detected in Southern blots by a series of V_H probes under a stringency that would detect 80% homology. Interfamily homologies, as determined by sequencing representative members, are 50–70% (39). As noted above, it has been shown that the adult expresses particular V_H genes or gene families roughly in proportion to their representation in the germline repertoire (18, 33, 34, 36).

The sample of IgM anti-IgGs described here includes six members of the J558 family, and one each from 7183 and J606 families (see references 39 and 43 for a description of V_H families). This would seem to show a preference for members of the J558 family; however, this family encompasses ~60–70% of the germline V_H genes (39, 43). Thus, finding J558 V_H genes in 75% of the IgM anti-IgG hybrids does not constitute a statistically significant preference for usage of members of this family. J606 and 7183 each constitute ~10% of the germline V genes; their presence in one of eight examples is not unexpected in a random selection model. For heavy chains, then, the null hypothesis of random selection cannot be rejected. It should be noted, however, that the two non-J558 V_H sequences occurred in hybridomas made from secondary protein, immunization-stimulated cell fusions. One other hybrid made in this way does contain a J558 heavy chain. Thus it is possible (but not significant with this sample size) that

non-J558 V_Hs are preferentially but not exclusively selected when RFs are generated by secondary protein immunization.

Other V Region Elements

As summarized in Table I, a variety of D_H genes representing each of the D_H families (44), as well as unidentified D sequences, are used in IgM anti-IgG1 monoclonals. Each of the J_Hs is used at least once. The combined D-J plus "N" segments (45) that include all of complementarity determining region (CDR) 3 are subject to length restrictions in certain kinds of antibody responses (18, 26, 46–48). Such a narrow restriction on CDR3 length does not apply to IgM anti-IgG, as in this set the third CDRs are of a variety of lengths, ranging from 7 to 10 residues. Shorter and longer CDR3s have been observed in other types of antibodies (49); it is possible that some restrictions on CDR3 length may apply to heavy chain variable segments found in IgM anti-IgG_1 autoantibodies. Finally, only $J_\kappa 1$ and $J_\kappa 2$ are represented in these sequences; this may reflect either a preference for these segments or the fact that $J_\kappa 1$ and $J_\kappa 2$ are used in 80% of splenic B cell V-J rearrangements (50).

Independent Monoclonal IgM Anti-IgG1 Autoantibodies Express Similar Light Chains in Combination with Dissimilar Heavy Chains

The data presented above show nonrandom selection of light chains and are consistent with random selection of heavy chains. The finding of instances in which the same or similar V_κs were paired with dissimilar V_Hs would further support the notion that there is little restriction on the usage of V_H in RFs. Fig. 3 shows two pairwise comparisons of heavy and light chain sequences from two sets of hybridomas that have similar or identical light chain sequences. In Fig. 3*a*, the light chains are from two different genes from the $V_\kappa 19$ family, and are 82% homologous at the amino acid level and 89% homologous at the nucleotide level. In contrast, the heavy chains, which are from the 7183 and J558 families, are 44% homologous at the amino acid level. A similar picture is presented by the sequences in Fig. 3*b*, in which the $V_\kappa 1$ light chains are identical except at the point of junctional diversity, whereas the heavy chain amino acid sequences are 42% homologous. In both of these pairs, different D_H families, different J_Hs, and different CDR3 lengths are used in the two heavy chains (see Table I). Thus, not only do heavy chains appear to be randomly represented in the set of IgM anti-IgG as a whole, but at least for $V_\kappa 19$ and $V_\kappa 1$ genes, very different heavy chains can associate with similar or identical light chains to generate the RF specificity.

There Is Very Little Somatic Mutation in IgM Anti-IgG1 Autoantibodies

Because RFs express only a small subset of V_κs, we have been able to observe three cases in which the same or nearly the same sequence is expressed in two independent IgM anti-IgG hybridomas. As outlined below, each of the three cases almost certainly represents the expression of the same germline gene in the pairs of hybridomas. By comparing the nucleotide sequences from each pair with each other and, when possible, with other published sequences, we have been able to scan a large tract of sequence for the occurence of somatic mutations.

RF49 and RF51. The nucleotide sequences of RF49 A and RF51 V_κ tran-

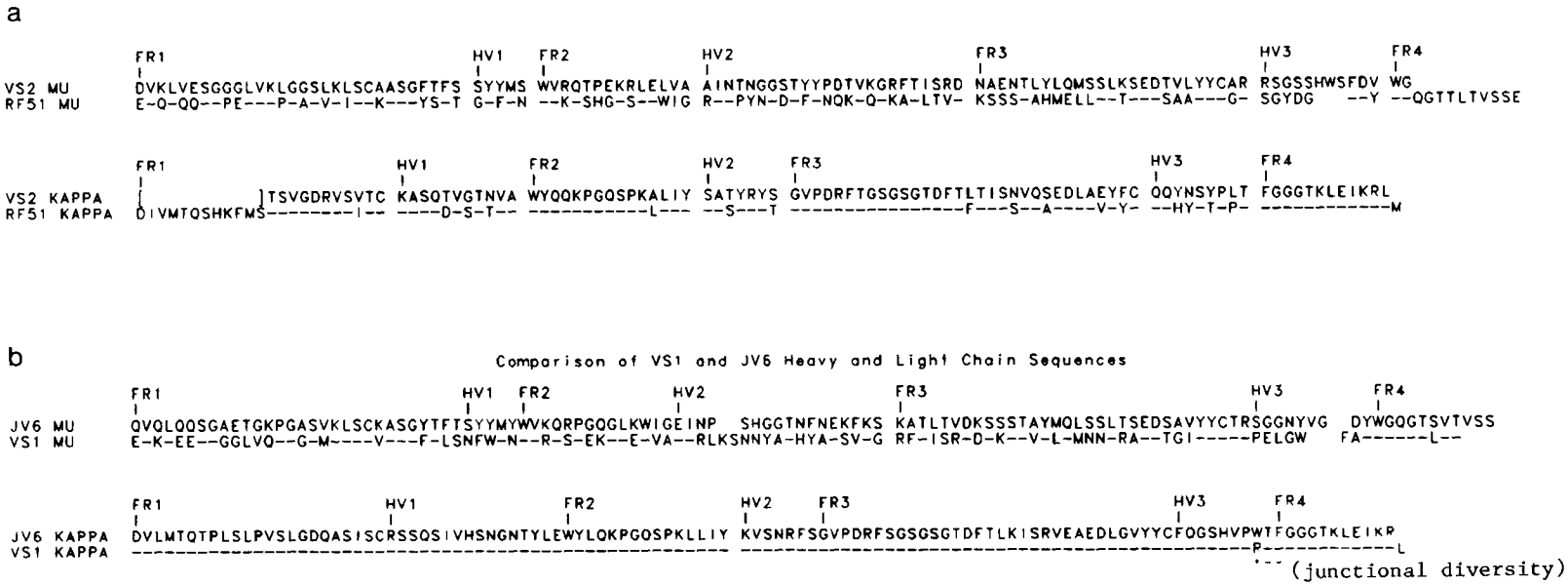

FIGURE 3. Comparison of heavy and light chain sequences from two pairs of IgM anti-IgG1 (RF) hybridomas. The sequences are derived and displayed as described in Fig. 1. (*a*) VS2 and RF51, which express different, 82% homologous genes from the V_x19 family; (*b*) VS1 and JV6, which express identical genes from the V_x1 family, differing only at the point of junction between V and J (indicated). In *a* the heavy chains are 44% homologous and in *b* they are 42% homologous.

scripts are identical. The amino acid sequences of both of these lines are also identical to the MPC11 amino acid sequence, including a small duplication in framework region (FR) 1 which is an idiosyncratic feature of the MPC11 gene (49, 51). We interpret the sequence identities to mean that RF49 and RF51 are expressing, in an unmutated form, the same germline gene as MPC11. Because the hybridomas have similar specificities, one must consider the otherwise remote possibility that the identity of the two sequences is the result of both having undergone identical parallel replacement mutations. However, this is unlikely to be the case for RF49 and RF51. Not only do the amino acid sequences match MPC11, an Ig of unknown specificity, but they also match the nucleotide sequence from another hybridoma of different specificity (M. Shlomchik, D. Pisetsky, and M. Weigert, unpublished results.)

VS1 and JV6. As mentioned above, the VS1 and JV6 hybrids also express identical nucleic acid sequences, except at the point of V-J joining. This is presumptive evidence that the two hybrids express the same germline gene and that the gene is unmutated in both cell lines. In particular, the amino acid sequence matches the consensus for the V_κ1C subgroup (27, 28), and therefore, both VS1 and JV6 probably express this germline gene. As with RF49 and RF51, identical mutations in both hybrids would result in the two having the same sequence. The possibility of a single parallel mutation can be ruled out by comparing this sequence with that of hybridomas of different specificity that are also expressing this gene. Hybridomas from the secondary responses to oxazolone (52) express the same V_κ1C gene. The consensus V_κ sequence from five independent anti-ox hybridomas is also identical to the VS1 and JV6 V_κ sequences. Thus, these chains contain no somatic mutations.

VS4 and JV2. The VS4 and JV2 lines have V_κ8 family sequences that differ by one nucleotide. The most likely interpretation of this is that they are both expressing the same gene and that one has incurred a somatic mutation. Two additional V_κ sequences of IgM anti-IgG2a hybridomas that are identical to JV2 (M. Shlomchik, J. Van Snick, and M. Weigert, unpublished data) support the idea that the VS4 V_κ contains a single mutation and the JV2 V_κ is unmutated.

Based on the above considerations, we have been able to survey ~2.0 kb of the V_κ region for somatic mutation and we have found one mutation. In addition, the J_κ and J_H region sequences in all of the hybrids are identical to the germline sequences. This represents another ~600 bases of unmutated sequences. Thus, the region that we have surveyed contains one mutation in 2.6 kb.

Discussion

We have found that a set of murine IgM anti-IgG1 monoclonal autoantibodies have light chains drawn from a small subset of V_κ genes and heavy chains drawn at random from a variety of V_H genes. This finding is analogous to results of analysis of human monoclonal RFs. Kunkel and coworkers originally showed (53, 54) that the majority of human monoclonal RFs have crossreactive idiotypes, and, by protein sequencing, that most RF light chains belong to the minor V_κ IIIb sub-subgroup. The finding of homologous V_H sequences in the Lay and Pom RF myeloma proteins, however, suggested a role for heavy chains in RF specificity (55). Subsequently, Chen et al. (56, 57), using antiidiotypic antibodies

against a synthetic peptide corresponding to regions of a human RF to screen a panel of monoclonal human RF proteins, showed that the light chains, but not the heavy chains, were the common element on human RF autoantibodies.

From these correlations of V_κ, but not V_H, we infer that anti-IgG specificity is determined largely or entirely by the light chain. We identified three recurrent families of V_κ sequences expressed in hybridomas; a significant percentage of B cells is expected to express a light chain gene derived from one of these families (see Results). Because our data imply that a variety of V_H, D_H, J_H, and J_κ seem competent for RF specificity, it follows that many of the B cells expressing RF V_κ family genes will also be RF-producing. Therefore, the V gene structure of RF mAbs provides an explanation for the puzzling observation that a high fraction of splenic B cells express IgM anti-IgG (3, 5).

Assumptions based on the types of V_κ, J_κ, V_H, D_H, and J_H found in RFs allow us to derive an estimate of the percentage of B cells containing the proper combination of V gene segments for the anti-IgG specificity. We have noted that 5–17% of B cells are expected to express apppropriate V_κ genes (see Results). J_κs in our sample are all $J_\kappa 1$ or $J_\kappa 2$; this could reflect a requirement for these segments or that 80% of B cells are known to use these two J_κs (50). In either case, 80–100% of B cells are expected to use an appropriate J_κ. We have observed V_H from three gene families which together encompass ~80% of all V_H genes. In addition, in two cases we found pairs of RFs that had nearly identical light chains in combination with very different heavy chains. This leads us to conclude that a variety of heavy chains, perhaps most, could be used for IgG1 binding. Our sample is too small to determine whether other V_H gene families are not represented because they are inappropriate or because of statistical fluctuation. We are currently expanding our analysis and carrying out other types of experiments, which will be necessary to delineate exactly what, if any, restrictions exist for V_H gene usage in RFs. Provisionally, we will nonetheless assume that 80% of B cells express a V_H competent for anti-IgG1 specificity. We have observed examples of each D_H family and all four of the J_Hs combined to create CDR3s of a variety of lengths. We take it that most D_H and J_H can be used in RFs (i.e., nearly 100% of cells express appropriate D_H and J_H).

Multiplying these percentages gives a value of 3–13% of virgin B cells expected to express V gene combinations encoding the anti-IgG specificity. This agrees remarkably well with the frequency observed in LPS and secondary protein immunization hybridomas (3–10%). A contrasting result of a similar estimate from a different antigen-antibody system is one derived by Manser, Huang, and Gefter (18) for the precursor frequency of the major Ars-CRI. The expression of this idiotype is thought to depend on single particular genes for each of the V region segments, and can thus be viewed as the opposite extreme of RF. Their estimate for precursor frequency is quite low: 1 Ars-CRI precursor per 4×10^7 B cells.

This idea provides a molecular explanation for a high precursor frequency of IgM anti-IgG1 splenic B cells. In principle, however, the relatively high frequency of such B cells could be the result of chronic stimulation by autoantigen and expansion of a small number of precursors. Several lines of evidence from previous studies, along with our present data, argue against this alternative: (*a*) cell mixing experiments in an in vivo adoptive transfer system have shown that

unprimed B cells are as effective as primed B cells in producing anti-IgG after immune complex injection or during a secondary response to a protein antigen; (*b*) T cells must be present to generate a response, and unprimed T cells are not as effective as primed T cells in providing the necessary help (6–8); (*c*) BALB/c *nu/nu* mice make a normal amount of IgM anti-IgG in response to LPS, but they do not make any IgM anti-IgG in response to immune complexes (indicating that their high frequency of LPS-driven cells is not the result of endogenous chronic stimulation, which is absent, as shown by the lack of response to immune complexes in these mice) (7); (*d*) all or nearly all anti-IgG activity is of the IgM isotype, contrary to what would be expected from a response based on chronically stimulated clones of B cells (3, 5, 58); and (*e*) the affinity of monovalent IgM anti-IgG for monovalent IgG is rather low, in contrast to what would be expected from an "affinity-matured" response to constant stimulation by antigen (52, 59). A correlative prediction of the chronic stimulation model is that the clonally expanded IgM anti-IgG should have a high content of somatic mutations (19, 20, 60, 61). In contrast, our data show an exceptionally low content of somatic mutations in IgM anti-IgG. Thus, we favor the model by which RFs can be constituted by a restricted set of light chains in combination with a variety of other variable region segments, thereby accounting for a high precursor frequency of RF-producing cells.

Recognizing that light chains of RFs might be important in determining specificity, we felt that protein sequence comparisons of IgM anti-IgG κ chains would provide an opportunity to discern regions that were shared among V_κ groups found in RFs, but not other V_κ sequences, and thus that might be important for conferring the ability to bind IgG1. In Fig. 1, the RF V_κ sequences are aligned with representative control sequences (presumably non-RF V_κs). The control sequences represent members of each of the V_κ families with the same CDR1 length as V_κ19, selected because these gene families are more related to V_κ19 sequences (which dominate the RF V_κ sequences) than are other V_κs (27).

Although, as expected, the RF light chain sequences within a group are quite homologous across the entire V region, we failed to find, in any of the three CDRs, greater self-similarity among the three RF V_κ families than among control V_κ families.[2] To our surprise, we noted instead that the FR2 and FR3 regions of all of the RF V_κ sequences are much more related to each other than are control sequences. This is evident by inspection of Fig. 1. The RF sequences differ from the consensus RF sequence by zero or one differences in FR2 (average, 0.7), the control sequences differ by 3–7 residues (average, 4.6); in FR3, the RFs differ from the consensus by 0–6 residues (average, 3.2), while the control sequences

[2] The null hypothesis under test is that particular regions (e.g., FR2) are no more homologous to each other among the RF V_κ groups than are non-RF V_κ groups. Disproving the null hypothesis is the criterion for establishing "significant homology." The test was performed by selecting at random single sequences representative of each V_κ family found in RFs, and as control to represent non-RF V_κs, single sequences representative of each V_κ family in length group 35 (as listed by Kabat et al., reference 47). All pairwise combinations of sequences within each of the two sets were generated and the number of differences between each pair was determined. The list of such values derived from all generated pairs for each set of V_κ sequences (i.e., RF or control) represents an assessment of self-similarity within that set. The control values presumably are a minimum assessment of expected self-similarity among all V_κs since the length group 35 V_κ families are evolutionarily more related to each other than are randomly selected V_κ families. The significance of differences between the two sets of values was then assessed by the Mann-Whitney test. The p values were: CDR1, not determined (because the RF CDR1s are of different lengths); CDR2, 0.14; and CDR3, 0.21.

differ from the consensus by 10–14 residues (average, 12.0). By the same analysis used to test CDR homology,[2] p values determined for FR2 and FR3 were 0.0045 for both. Moreover, a survey of all other V_κ groups not shown in Fig. 1 reveals that only one other known V_κ group ($V_\kappa 22$) has a sequence that resembles $V_\kappa 1$, 8, and 19 across FR2 and FR3.

The FR2 and FR3 homology in RF V_κs is even more striking when one considers that the $V_\kappa 1$, $V_\kappa 8$, and $V_\kappa 19$ gene families are divergent in other regions of their sequence. As can be seen in Fig. 1, each group has a different CDR1 length and is quite different in sequence from the other two through FR1 and all CDRs. A dendrogram of V_κ relatedness based on protein sequence to the first invariant tryptophan residue shows $V_\kappa 19$ diverging from $V_\kappa 8$ and $V_\kappa 1$ at the very first node; $V_\kappa 1$ and $V_\kappa 8$ in turn diverge at the second node (27). Similarity of FR2 and FR3 in RF V_κ sequences leads us to the unorthodox proposal that these regions include the combining site for IgG1. A corollary of this hypothesis is that all of the RF we have sequenced make similar contacts with an Ig constant domain. Consistent with this is the result that most, but not all, murine RF bind the CH3 domain of IgG (62).

An important test for a model that postulates antigen contact with FR2 and FR3 is whether any segments of these regions are on the surface of the molecule, therefore accessible for interaction with other molecules. The most appropriate known three-dimensional structure for approaching this question is that of MOPC603 (63), which has a $V_\kappa 8$ light chain (27) and is very related in FR2 and FR3 to the RF sequences. As seen in Fig. 4, the protein-chain backbones of FR2 and FR3 trace paths that are almost entirely on the surface of the molecule; many of the R groups (not shown in Fig. 4) project out and are accessible. In particular, the central portion of FR3 is an antiparallel β sheet that forms a large flat surface (Fig. 4*a*). Contiguous with this surface, but in another plane, is an accessible region defined by FR2 and some of the *N*-terminal and *C*-terminal residues of FR3 (Fig. 4*b*). While we have no direct evidence for usage of these regions in binding, it is clear from their location in the molecule that they could be used for this purpose. Another indication that FR determinants are accessible is the finding of antiidiotypic sera that have been shown to bind FR determinants (64).

An alternative model for the interaction of IgM anti-IgG with antigen states that each of the three RF family V_κs uses the traditional antigen-combining site to bind a different epitope on IgG-CH3. Similarity of CDRs within each of the RF V_κ families (see Fig. 1) is the basis for this hypothesis. The strong FR2/3 homology, of the RF V_κ families is, in this model, ascribed to coincidence. This is a difficulty, because there are only 4 out of 24 known families that have such FR2/3 sequences, and the RF families happen to have these sequences in all three cases. Thus, the probability of this type of FR sequence randomly occurring in the three families that bind these hypothetically independent epitopes is no more than $(1/6)^3 = (1/216)$. Another problem for this model is that the light chain is restricted and the heavy chain is variable in antibodies that bind to each of the three postulated epitopes. This characteristic is unusual in other antigen-antibody systems described in the literature. In most cases, only certain combinations of heavy and light chains will allow binding of a given epitope (33, 48, 65–67a) although in a few cases light chain restriction with some degree of heavy

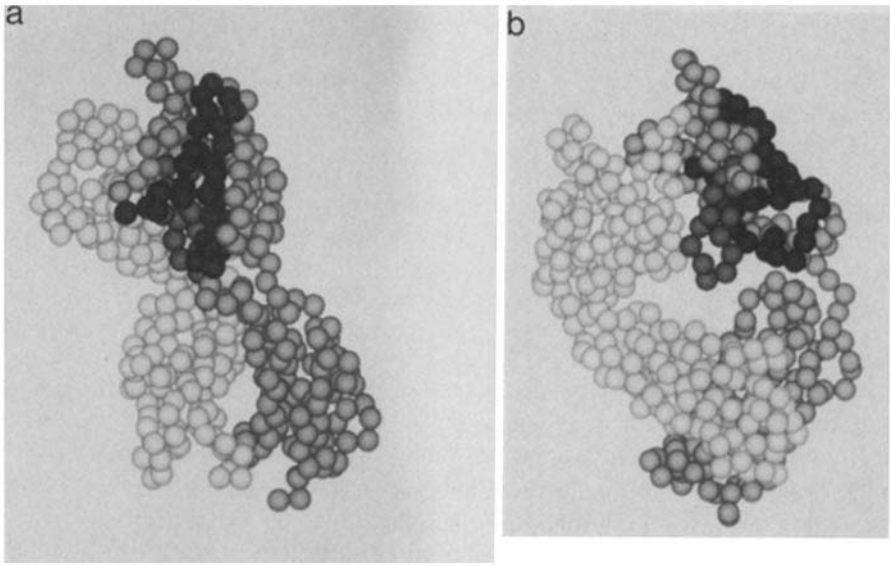

FIGURE 4. Two views of a computer-generated space-filling model of the α carbons of MCPC603 Fab (reference 63). Refined coordinates kindly provided by Dr. E. Padlan were used. Shading of spheres indicates: light, heavy chain; grey, light chain; dark grey, light chain FR2; black, light chain FR3. The figures are oriented so that V region domains are at the top and C domains at the bottom. (*a*) Highlights the large, exposed anti-parallel β sheet and bend of V_κ FR3; (*b*) highlights the fact that most of the V_κ FR2 backbone is exposed and that much of the *N*- and *C*-terminal portions of FR3 are similarly exposed, coplanar with the FR2 residues. The images were produced by first, interactive manipulation of the molecule on the Evans-Sutherland PS-300 using the DOCK graphics package, developed at the Fox Chase Cancer Center by W. Wood, F. Manion, and R. Stodola, followed by generation of a space-filling raster display of selected views using MOLDISP, by R. Stodola, for display on the AED512 raster graphics terminal.

chain variation (19, 65, 67) has been found. It seems unlikely that the same rather unusual heavy chain promiscuity should apply to all three epitopes. Because of these two substantial problems, we do not favor this alternative model, although the present data does not rule it out.

In this connection, we do not believe that the model we favor is contradicted by recent studies of the human system by Chen et al. (56), which showed that an anti-CDR2-peptide serum could detect most human RF V_κs. This correlation seems to implicate the CDR2 structure in binding of antigen. However, it is reasonable to assume that the population of molecules expressing the CDR2 sequences detected by the rabbit antisera would also be expressing homologous sets of FR regions. In fact, a concordance of CDR and FR sequences is evidenced in the known human RF sequences (see reference 49). An analogous anti-CDR2-peptide experiment done in the murine system using a V_κ19 CDR2 peptide would give a result similar to that found in the human system, although we know from the data presented here that nearly all, if not all, murine RF share homologous FR2 and FR3 sequences.

If, as our results suggest, RF binds an Ig constant domain through an unusual combining site (FR2/3), then RF can be viewed as antibodies that possess two functional combining sites, one for IgG-CH3 made up of FR2/3 residues and

the other of unknown and variable specificity that comprises V_H-V_1 CDR residues. Since RF are made by a relatively large subset of B cells in which light chains are derived from three families with dissimilar CDRs and in which heavy chains are diverse, a large CDR repertoire is probably available in this subset.

Regarding this, it is notable that some monoclonal RF have been found to have dual specificities (68–71, and D. Nemazee, unpublished observation). Two interesting examples of polyspecificity are RF that bind β-2-microglobulin (including VS1 and VS2, D. Nemazee, unpublished data) and a class of antibody that has specificity for both an Ig-V region (i.e., they are antiidiotype) and for the Fc portion of IgG (70, 71) (i.e., they are RF). We hypothesize that in this latter class of antibody (termed either "epibodies" or "homobodies," depending on the network of interactions that induced them), the conventional binding site interacts with the idiotope because the CDRs of $V_\kappa 1$, $V_\kappa 8$, or $V_\kappa 19$ happen to be suitable for this purpose, while the FR2 and FR3 of the light chains of these antibodies mediate IgG binding.

Dual specificity need not always be mediated through two different combining sites. An alternative and intriguing explanation for crossspecificity for β-2-microglobulin is that β-2-microglobulin and the structure recognized by RF are both Ig family domains (72) and probably have similar shapes. It is possible then that recognition of β-2-microglobulin could be via the same combining site that recognizes IgG-Fc: FR2 and FR3. This is in fact suggested by the complete divergence of VS1 and VS2 V_H and V_κ sequences in all regions except for V_κ FR2 and FR3 (see Figs. 1 and 2).

The VS1 antibody is particularly interesting in this context. It is exceptional in that it binds a mutant IgG1 (IF1) that lacks the CH3 domain; therefore, VS1 may be recognizing the CH2 domain (D. Nemazee, unpublished data). Since its light chain type, $V_\kappa 1$, is prevalent among RF (this study and M. Shlomchik, J. Van Snick, D. Pisetsky, and M. Weigert, unpublished data), and since, in one investigation, 70 of 71 monoclonal RF failed to bind the truncated IgG1 mutant IF1 (62), it is probable that many $V_\kappa 1$-type RF other than VS1 bind CH3, not CH2. Moreover, we have sequenced an anti-IgG2a RF which in fact binds CH3, and we found it to be expressing $V_\kappa 1$ (M. Shlomchik, J. Van Snick, and M. Weigert, unpublished data). These considerations raise the possibility that IgG-CH2 and IgG-CH3 are crossreactive for some RF.

In a formal sense, the recognition of IgG by RF represents an interaction between Ig family domains. It is interesting to speculate that the interaction of two other Ig family molecules, the T cell receptor and MHC, might be structurally analogous. In particular, the regions of the T cell receptor β chain that are homologous to κ FR2 and FR3 might be used to bind one of the MHC domains. The increased variability reported among β chain sequences in these locations (73) in turn may reflect the need to bind a variety of different MHC alleles (whereas constancy in these regions for V_κ reflects evolution to bind a single, conserved structure: IgG-CH3). Since all functional T cell receptors must recognize MHC (74), and since the same V_β gene can be used to bind class I and class II molecules (75, 76), a generic type of binding based on this kind of interaction between Ig family domains is an attractive possibility for a component of MHC recognition. Such a model is consistent with a recent report by Yague et al. (77), in which they characterized a variant of a chicken OVA (cOVA) +

I-A^k and I-A^b-reactive T cell hybridoma that has lost specific α chain and reactivity to cOVA + I-A^k, but retains alloreactivity to I-A^b.

The nature of the IgM anti-IgG sequences that we have presented here has led to several hypotheses about the heavy and light chain composition of RF, the way in which RF interacts with IgG, and the way Ig family domains may recognize each other. Although some of these ideas are unconventional, they are testable and have suggested appropriate chain recombination experiments to explore the V region structural requirements for IgG binding, as well as biochemical and structural studies to determine how RF does indeed bind IgG.

Summary

The nucleotide sequences of heavy and light chains from 10 monoclonal IgM anti-IgG1 (RF) antibodies were determined and reported here as translated amino acid sequences. Only three families of V_κ light chains were used in these antibodies: $V_\kappa 1$ (two examples), $V_\kappa 8$ (three examples), and $V_\kappa 19$ (four examples). This represents a significant nonrandom selection of light chains. In contrast, all other variable region gene segments (i.e., V_H, D_H, J_H, and J_κ) were used in a pattern consistent with random selection from the available pool of germline genes. In two cases, the same anti-IgG1 specificity was generated by a combination of very homologous light chains with unrelated heavy chains. We infer from this that the light chain is the segment used by these antibodies to bind IgG1.

The nature of these sequences provides an explanation for the curious observation that as many as 15% of splenic B cells in normal mice may be expressing IgM anti-IgG; if, as our data suggest, certain light chains in combination with many different heavy chains can be used in assembling the anti-IgG specificity, then, because of combinatorial association in which the heavy chain is not relevant for specificity, the fraction of IgM-producing B cells expressing these light chains should approximate the fraction of B cells making IgM anti-IgG. We calculate, based on data presented in several other studies, that 5–17% of B cells express one of the V_κ types observed in monoclonal RF. This agrees well with estimates for the number of B cells making IgM anti-IgG. In addition, our findings could rule out other explanations of the high percentage of B cells making RF, such as constant stimulation by antigen or presence of numerous antigenic epitopes since it was shown that IgM anti-IgG1 antibodies are not somatically mutated and that they are structurally homogeneous.

We aligned the V_κ sequences of the RF in hopes of finding some primary sequence homology between the represented V_κ families which might point to residues involved in the binding interaction. Although we found no such homology in the hypervariable regions, we did find significant and unexpected homology in the FR2 and FR3 of these light chains. We noted that these regions are exposed in the Ig structure and postulate that they may be involved in a unique type of binding interaction between two Ig family domains, i.e., V_κ binding to a constant region domain of IgG.

We thank Dr. Andy Caton for providing the initial protocol upon which our sequencing technique is based; Dr. Sam Litwin for extensive discussions and statistical analysis; Dr. Helen Berman for help with computer graphics; Dr. Eduardo Padlan for discussions and providing coordinates for MCPC603; Ms. Lynda Peitzman for expert technical help; and

Mr. E. Paul Kaplan, a summer undergraduate student, for performing several excellent sequence determinations.

Received for publication 2 December 1985 and in revised form 3 April 1986.

References

1. Izui, S., R. Eisenberg, and F. Dixon. 1979. IgM rheumatoid factors in mice injected with bacterial LPS. *J. Immunol.* 122:2096.
2. Dresser, D. 1978. Most IgM-producing cells in the mouse secrete auto-antibodies (rheumatoid factor). *Nature (Lond.).* 274:480.
3. Van Snick, J., and P. Coulie. 1982. Monoclonal Anti-IgG autoantibodies from lipopolysaccharide-activated spleen cells of 129/Sv mice. *J. Exp. Med.* 155:219.
4. Nemazee, D., and V. Sato. 1983. Induction of rheumatoid antibodies in the mouse. Regulated production of autoantibody in the secondary humoral response. *J. Exp. Med.* 158:529.
5. Van Snick, J., and P. Coulie. 1983. Rheumatoid factors and secondary immune responses in the mouse. I. Frequent occurrence of hybridomas secreting IgM anti-IgG1 autoantibodies after immunization with protein antigens. *Eur. J. Immunol.* 13:890.
6. Coulie, P., and J. Van Snick. 1983 Rheumatoid factors and secondary immune responses in the mouse. II. Incidence, kinetics, and induction mechanisms. *Eur. J. Immunol.* 13:895.
7. Nemazee, D. 1985. Immune complexes can trigger specific, T cell-dependent, autoanti-IgG antibody production in mice. *J. Exp. Med.* 161:242.
8. Coulie, P. and J. Van Snick. 1985. Rheumatoid Factor (RF) production during anamnestic immune responses in the mouse. III. Activation of RF precursor cells is induced by their interaction with immune complexes and carrier-specific helper T cells. *J. Exp. Med.* 161:88.
9. Welch, M., S. Fong, J. Vaughan, and D. Carson. 1983. Increased frequency of rheumatoid factor precursor B lymphocytes after immunization of normal adults with tetanus toxoid. *Clin. Exp. Immunol.* 51:299.
10. Abruzzo, J., and C. Christian. 1961. Induction of a rheumatoid factor-like substance in rabbits. *J. Exp. Med.* 114:791.
11. Clarkson, A., and G. Mellow. 1981. Rheumatoid Factor-like immunoglobulin M protects previously uninfected rat pups and dams from *Trypanosoma lewisi. Science (Wash. DC).* 214:186.
12. Rose, H., C. Ragan, E. Pearie, and M. Lipman. 1948. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoidd arthritis. *Proc. Soc. Exp. Biol. Med.* 68:1.
13. Andersson, J., A. Coutinho, and F. Melchers. 1977. Frequencies of mitogen-reactive B cells in the mouse II. Frequencies of B cells producing antibodies which lyse sheep or horse erthrocytes, and trinitrophenylated or nitroiodophenylated sheep erythrocytes. *J. Exp. Med.* 145:1520.
14. Levy, M. 1984. Frequencies of phosphorylcholine-specific and T15-associated 10/13 idiotope-positive B cells within lipopolysaccharide-reactive B cells of adult BALB/c mice. *Eur. J. Immunol.* 14:864.
15. Augustin, A., and A. Coutinho. 1980. Specific T helper cells that activate B cells polyclonally: in vitro enrichment and cooperative function. *J. Exp. Med.* 151:587.
16. Bernabe, R., A. Coutinho, C. Martinez-A., and P. Cazenave. 1981. Immune networks: frequencies of antibody- and idiotype-producing B cell clones in various steady states. *J. Exp. Med.* 154:552.

17. Primi, D., F. Mami, C. Le Guern, and P. Cazenave. 1982. Mitogen-reactive B cell subpopulations selectively express different sets of V regions. *J. Exp. Med.* 156:181.
18. Manser, T., S. Huang, and M. Gefter. 1984. Influence of clonal selection on the expression of immunoglobulin variable region genes. *Science (Wash. DC).* 226:1283.
19. Clarke, S., K. Huppi, D. Ruezinsky, L. Staudt, W. Gerhard, and M. Weigert. 1985. Inter- and intraclonal diversity in the antibody response to influenza hemagglutinin. *J. Exp. Med.* 161:687.
20. Sablitzky, F., G. Wildner, and K. Rajewsky. 1985. Somatic mutation and clonal expansion of B cells in an antigen-driven immune response. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:345.
21. Chirgwin, J., A. Przybyla, R. MacDonald, and W. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.
22. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratories, Cold Spring Harbor, New York. 545 pp.
23. Rubin, C., and C. Schmid. 1980. Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res.* 8:4613.
24. Bencini, D. G. O'Donovan, and J. Wild. 1984. Rapid chemical degradation sequencing. *Biotechniques.* Jan./Feb.:p. 4.
25. Biggin, M., J. Gibson, and G. Hong. 1983. Buffer gradient gels and ^{35}S label as aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA.* 80:3963.
26. Kaartinen, M., G. Griffiths, A. Markham, and C. Milstein. 1983. mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone and its early diversification. *Nature (Lond.).* 304:320.
27. Potter, M., J. B. Newell, S. Rudikoff, and E. Haber. 1982. Classification of mouse V kappa groups based on the partial amino acid sequence to the first invariant tryptophan: impact of 14 new sequences from IgG myeloma proteins. *Mol. Immunol.* 19:1619.
28. Gibson, D. M. 1984. Evidence for 65 electrophoretically distinct groups of light chains in BALB/c and NZB myelomas. *Mol. Immunol.* 21:421.
29. Lejeune, J. L., D. E. Briles, A. R. Lawton and J. F. Kearney. 1982. Estimate of the light chain repertoire size of fetal and adult BALB/cJ and CBA/J mice. *J. Immunol.* 129:673.
30. Julius, M., D. Mckean, M. Potter, and M. Weigert. 1981. Expression of kappa chains of the $V_\kappa 21$ group in *Mus musculus* and related species. *Mol. Immunol.* 18:11.
31. Cory, S., B. Tyler, and J. Adams. 1981. Sets of immunoglobulin V kappa genes homologous to ten cloned V kappa sequences: implication for the number of germline V kappa genes. *J. Mol. and Appl. Genet.* 1:103.
32. Moynet, D., S. Maclean, K. Ng, D. Anctil, and D. Gibson. 1985. Polymorphism of kappa-variable region ($V_\kappa 1$) in inbred mice: relationship to the Igk-Ef2 serum light chain marker. *J. Immunol.* 134:3455.
33. Near, R., T. Manser, and M. Gefter. 1985. The generation of major and minor idiotype-bearing families of anti-*p*-azophenylarsonate antibodies: stochastic utilization of V_h gene segments. *J. Immunol.* 134:2004.
34. Perlmutter, R., J. Kearney, S. Chang, and L. Hood. 1985. Developmentally controlled expression of immunoglobulin V_H genes. *Science (Wash. DC).* 227:1597.
35. Weigert, M., and R. Riblet. 1976. Genetic control of antibody variable regions. *Cold Spring Harbor Symp. Quant. Biol.* 41:837.
36. Basta, P., H. Kubagawa, J. Kearney, and D. Briles. 1983. Ten percent of normal B cells and plasma cells share a V_h determinant(s) (J606-GAC) with a distinct subset of murine V_h III plasmacytomas. *J. Immunol.* 130:2423.

37. Yancopoulos, G., S. Desiderio, M. Paskind, J. Kearney, D. Baltimore, and F. Alt. 1984. Preferential utilization of the most J_h-proximal V_h gene segments in pre-B-cell lines. *Nature* (*Lond.*). 311:727.
38. Hawley, R., M. Shulman, H. Murialdo, D. Gibson, and N. Hozumi. 1982. Mutant immunoglobulin genes have repetitive DNA elements inserted into their intervening sequences. *Proc. Natl. Acad. Sci. USA.* 79:7425.
39. Brodeur, P., and R. Riblet. 1984. The immunoglobulin heavy chain variable region in the mouse I. 100 IgH-V genes comprise 7 families of homologous genes. *Eur. J. Immunol.* 14:922.
40. Crews, S., 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 the antibody. *Cell.* 25:59.
41. Loh, D., A. Bothweel, M. White-Scharf, T. Imanishi-Kari, and D. Baltimore. 1983. Molecular basis of a mouse strain-specific anti-hapten response. *Cell.* 33:85.
42. Joho, R., and I. Weissman. 1984. Evolution of a multigene family of V_κ germline genes. *EMBO* (*Eur. Mol. Biol. Organ.*) *J.* 3:185.
43. Brodeur, P., M. Thompson, and R. Riblet. 1984. The content and organization of mouse IgH-V families. *In* Regulation of the Immune System. H. Cantor, L. Chess, and E. Sercarz, editors. Alan R. Liss, Inc., New York. 445.
44. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *J. Exp. Med.* 155:201.
45. Alt, F., 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.
46. Wysocki, L., M. Margolies, B. Huang, D. Nemazee, D. Wechsler, V. Sato, J. Smith, and M. Gefter. 1985. Combinational diversity within variable regions bearing the predominant anti-*p*-azophenylarsonate idiotype of strain A mice. *J. Immunol.* 134:2740.
47. Rudikoff, S., M. Pawlita, J. Pumphrey, E. Mushinsky, and M. Potter. 1983. Galactan-binding antibodies: diversity and structure of idiotypes. *J. Exp. Med.* 158:1385.
48. Darsley, M., and A. Rees. 1985. Nucleotide sequences of five anti-lysozyme monoclonal antibodies. *EMBO* (*Eur. Mol. Biol. Organ.*) *J.* 4:393.
49. Kabat, E., T. Wu, H. Bilofsky, M. Reid-Miller, and H. Perry. 1983. Sequences of proteins of immunological interest. U. S. Government Printing Office, Bethesda, MD.
50. Wood, D., and C. Coleclough. 1984. Different joining region J-elements of the murine kappa-immunoglobulin locus are used at markedly different frequencies. *Proc. Natl. Acad. Sci. USA.* 81:4756.
51. Smith, G. 1978. Sequence of the full length immunoglobulin kappa chain of mouse myeloma MPC11. *Biochem. J.* 171:337.
52. Berek, C., G. Griffiths, and C. Milstein. 1985. Molecular events during maturation of the immune response to oxazolone. *Nature* (*Lond.*). 316:412.
53. Kunkel, H., V. Agnello, F. Joslin, R. Winchester, and J. Capra. 1973. Cross-idiotypic specificity among monoclonal IgM proteins with anti-gamma-globulin activity. *J. Exp. Med.* 137:331.
54. Pons-Estel, B., F. Goñi, A. Solomom, and B. Frangione. 1984. Sequence similarities among κIIIb chains of monoclonal human IgMκ autoantibodies. *J. Exp. Med.* 160:893.
55. Capra, J., and J. Kehoe. 1974. Structure of antibodies with shared idiotypy. The complete sequence of the heavy chain variable regions of two immunoglobulin M anti-gamma globulins. *Proc. Natl. Acad. Sci. USA.* 71:4032.
56. Chen, P., F. Goni, S. Fong, F. Jirik, J. Vaughan, B. Frangione, and D. Carson. 1985.

The majority of human monoclonal IgM rheumatoid factors express a "primary structure-dependent" crossreactive idiotype. *J. Immunol.* 134:3281.

57. Chen, P., F. Goni, R. Houghten, S. Fong, R. Goldfein, J. Vaughan, B. Frangione, and D. Carson. 1985. Characterization of human rheumatoid factors with seven anti-idiotypes induced by synthetic hypervariable region peptides. *J. Exp. Med.* 162:487.
58. Van Snick, J., and P. Masson. 1980. Incidence and specificities of IgA and IgM anti-IgG autoantibodies in various mouse strains and colonies. *J. Exp. Med.* 151:45.
59. Eisenberg, R. 1976. The specificity and polyvalency of binding of a monoclonal rheumatoid factor. *Immunochemistry.* 13:355.
60. Reth, M., G. Hammerling, and K. Rajewski. 1978. Analysis of the repertoire of anti-NP antibodies in C57B1/6 mice by cell fusion I. Characterization of antibody families in the primary and hyperimmune response. *Eur. J. Immunol.* 8:393.
61. Gearhart, P., N. Johnson, R. Douglas, and L. Hood. 1981. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature (Lond.).* 291:29.
62. Stassin, V., P. Coulie, B. Birshtein, D. Secher, and J. Van Snick. 1983. Determinants recognized by murine rheumatoid factors: molecular localization using a panel of mouse myeloma variant immunoglobulins. *J. Exp. Med.* 158:1763.
63. Segal, D., E. Padlan, G. Cohen, S. Rudikoff, M. Potter, and D. Davies. 1974. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. USA.* 71:4298.
64. Julius, M., D. Mckean, M. Potter, R. Feldmann, and M. Weigert. 1981. The structural basis of antigenic determinants on $V_\kappa 21$ light chains. *Mol. Immunol.* 18:1.
65. Devaux, C., D. Moinier, G. Mazza, X. Guo, S. Marchetto, M. Fougereau, and M. Pierres. 1985. Preferential expression of $V_\kappa 21E$ on IdX Ia.7 positive monoclonal anti-I-E antibodies. *J. Immunol.* 134:4024.
66. Tonnelle, J. Roca-Serra, A. Moulin, D. Moinier, and M. Fougereau. 1983. V_κ gene family in (GAT)-specific antibodies that express CGAT (or pGAT) public idiotype specificities. *J. Exp. Med.* 158:1415.
67. Perlmutter, R., J. Klotz, M. Bond, M. Nahm, J. Davie, and L. Hood. 1984. Multiple V_H gene segments encode murine antistreptococcal antibodies. *J. Exp. Med.* 159:179.

67a. Caton, A., G. Brownlee, L. Staudt, and W. Gearhardt. 1986. Structural and functional implications of a restricted antibody response to a defined antigenic region on the influenza virus hemagglutinen. *EMBO (Eur. Mol. Biol. Organ.) J.* In Press.

68. Agnello, A. Arbetter, G. Ibanez de Kasep, R. Powell, E. Tan, and F. Joslin. 1980. Evidence for a subset of rheumatoid factors that cross-react with DNA-histone and have a distinct cross-idiotype. *J. Exp. Med.* 151:1514.
69. Rubin, R., R. Balderas, E. Tan, F. Dixon, and A. Theofilopoulos. 1984. Multiple autoantigen binding capabilities of mouse monoclonal antibodies selected for rheumatoid factor activity. *J. Exp. Med.* 159:1429.
70. Bona, C., S. Finley, S. Waters, and H. Kunkel. 1982. Anti-immunoglobulin antibodies III. Properties of sequential anti-idiotypic antibodies to heterologous anti-γ globulins. Detection of reactivity of anti-idiotype antibodies with epitopes of Fc fragments (homobodies) and with epitopes and idiotopes (epibodies). *J. Exp. Med.* 156:986.
71. Bona, C., B. Goldberg, D. Metzger, J. Urbain, and H. Kunkel. 1984. Anti-immunoglobulin antibodies IV. Cross-reaction of anti-idiotypic antibodies specific for rabbit and murine anti-a1 allotype antibodies with Fc fragment of human immunoglobulins. *Eur. J. Immunol.* 14:548.
72. Peterson, P., B. Cunningham, I. Berggard, and G. Edelman. 1972. Beta-2-microglobulin—A free immunoglobulin domain. *Proc. Natl. Acad. Sci. USA.* 69:1697.
73. Patten, P., T. Yokota, J. Rothbard, Y. Chien, K. Arai, and M. Davis. 1984. Structure,

expression and divergence of T-cell receptor beta-chain variable regions. *Nature* (*Lond.*). 312:40.

74. Heber-Katz, E., R. Schwartz, L. Matis, C. Hannum, T. Fairwell, E. Appella, and D. Hansburg. 1982. Contribution of antigen-presenting cell major histocompatibility complex gene products to the specificity of antigen-induced T cell activation. *J. Exp. Med.* 155:1086.
75. Rupp, F., H. Acha-Orbea, H. Hengartner, R. Zinkernagel, and R. Joho. 1985. Identical V_β T-cell receptor genens used in alloreactive cytotoxic and antigen plus I-A specific helper T cells. *Nature* (*Lond.*). 315:425.
76. Acuto, O., R. Hussey, and E. Reinherz. 1985. Multiple class I and class II major histocompatibility complex allospecificities are generated with T cell receptor variable (*V*) domains created by a single Ti β *V* gene family. *J. Exp. Med.* 162:1387.
77. Yague, J., J. White, C. Coleclough, J. Kappler, E. Palmer, and P. Marrack. 1985. The T cell receptor: the alpha and beta chains define idiotype, and antigen and MHC specificity. *Cell.* 42:81.