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

II. Comparison of Hybridomas Derived by Lipopolysaccharide Stimulation and Secondary Protein Immunization

BY MARK SHLOMCHIK,* DAVID NEMAZEE,‡ JACQUES VAN SNICK,§ AND MARTIN WEIGERT*

*From the *Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111; the ‡Basel Institute for Immunology, Postfach CH-4005, Basel, Switzerland; and the §Ludwig Institute for Cancer Research, Brussels, Belgium*

IgM autoantibodies that bind the constant region of IgG (IgM anti-IgG) are a normal physiological component of secondary immune responses in a number of species, including mouse and human (1–3). Such antibodies have been termed rheumatoid factors (RF),[1] owing to their original identification in the serum of rheumatoid arthritis patients, who constitutively produce RF (4). In the mouse, it has been shown that immune complexes formed during the course of a secondary immune response to a variety of protein antigens can stimulate normal B cells to produce IgM anti-IgG in a T helper cell–dependent manner (5, 6). In a secondary immune response to protein antigens, the quantity of IgM anti-IgG can equal or exceed the amount of antigen-specific antibody (2).

The fraction of B cells that can produce this autoantibody may also be very high; 3–10% of hybridomas derived from mice stimulated with the nonspecific B cell mitogen bacterial lipopolysaccharide (LPS) produce IgM that bind IgG (7). Our previous analysis of the variable region sequence of nine such antibodies suggested that a large number of different germline V_H genes, in association with a variety of D_H, J_H, and J_κ could be used in combination with a limited set of V_κ genes to encode the RF specificity (8). We proposed that this pattern of V gene expression could account for the high frequency of B cells that can produce IgM that binds IgG.

It has been suggested that the large fraction of LPS-activated B cells that make IgM anti-IgG is the reason for the large quantity of RF secreted during the secondary immune response (5, 6). However, an alternative model is that only a subset of the cells found among LPS-derived IgG-binding IgM antibodies are stimulated to produce RF in vivo during a secondary immune response. A

This work was supported by grant GM-20964-12 from the National Institutes of Health to M. Weigert, and by an appropriation from the Commonwealth of Pennsylvania. Mark 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; FR, framework region; RF, rheumatoid factor.

 J. EXP. MED. © The Rockefeller University Press · 0022-1007/87/04/0970/18 $1.00
Volume 165 April 1987 970–987

prediction of the first model is that the V regions expressed in the two types of RF should be similar. To allow a significant comparison of the V region structures of these two types of RF, we have analyzed 15 additional antibodies, providing a database of 10 LPS- and 14 secondary protein immunization–derived RFs. In this sample, the variable regions of IgG-binding IgM antibodies isolated after LPS stimulation are similar from those of RFs derived after secondary immunization with protein. This supports the idea that the large quantity of RF produced during the secondary immune response is attributable to a large number of B cells that can make this antibody. In addition, the new sequence data confirm and extend our earlier findings concerning V region usage in RFs, and provide further details of the possible means by which RF antibodies bind IgG.

Materials and Methods

Hybridomas. The VS and RF series of hybridomas were described previously, as were JV2 and JV6 (8). The remaining JV series hybridomas were produced by fusion with SP2/0 and tested by ELISA and particle agglutination assays as described (3). The A and B series hybridomas were produced by fusion with the nonsecreting cell line PAIO (M. Brockhaus, personal communication) and tested by hemolysis assay and by RIA that detected bound IgM. All hybridomas bind IgG1; some hybridomas bind other murine isotypes and some types of heterologous IgG, as is commonly found among IgM anti-IgG1 (3, and D. Nemazee and J. Van Snick, unpublished observation). Hybridomas were derived from spleens of mice stimulated either by in vivo LPS injection or by secondary immunization with antigenic protein (Table I). With the exception of the A series of hybridomas, RF hybridomas sequenced here or reported elsewhere react with the immunizing protein. The A series hybridomas did not react with arsonate-coupled red blood cells (Ars-RBC), but have not been assayed on Limulus polyphemus hemocyanin–Ars (LPH)-Ars. Spleen cells were derived from BALB/c mice (IgH^a) except for JV3-JV5 and JV8-JV10, which were from 129/Sv mice (IgH^a).

Sequence Determinations of Variable Regions of Ig mRNA. Growth of cells, preparation of RNA, production of labelled cDNA, and chemical degradative sequencing reactions were performed as described (8). Briefly, we used poly(A)-enriched RNA from hybridoma cells as a template for synthesis of full-length cDNA primed by 5′-labelled oligonucleotides that hybridized to the constant portions of either the κ (C_κ primer) or μ (C_μ primer) mRNAs at a site just 3′ of the V-C junction. The full-length cDNA was isolated on a polyacrylamide gel, eluted, and subjected to chemical degradative sequencing reactions. The resolution of the labelled cDNA fragments generated by these reactions on three types of polyacrylamide sequencing gels usually resulted in the unambiguous assignment of virtually all nucleotides up to ~400 bases from the priming site.

Results

15 IgM anti-IgG1 hybridomas were selected at random from RF hybridoma libraries assembled from several fusions performed in two laboratories. These are described in Table I. RNA from these cell lines was prepared and sequenced as described in Materials and Methods. The nucleotide sequences were read from autoradiograms of three overlapping gels and were translated to amino acid sequences using the universal genetic code.

Fig. 1 shows the amino acid sequences of the κ light chains of the 15 hybridomas along with similarly derived data from 9 hybridomas described earlier. Fig. 2 shows the heavy chain sequences grouped according to the type of light chain with which each is paired (see Discussion). For reference, the corresponding

TABLE I

Summary of Sequence Characteristics of IgM Anti-IgG1 Sequences

Cell	Derivation*	V_κ group‡	J_κ§	Vh family	D_H‖	J_H	CDR3 length¶
JV3	LPS	1d	2	J558	FL16.2	4	7
JV2	LPS/$DxSO_4$	8a	1	J558	NA	2	7
JV4	LPS	8a	1	J558	SP2.3,.4,.6	4	7
JV5	LPS	8a	4	J558	NA	2	7
VS3	LPS	8b	2	J558	NA	4	8
VS4	LPS	8a	1	J558	FL16.1,.2	3	7
RF34	LPS	19c	2	J558	SP2 (any)	4	8
RF49	LPS	19a	2	J558	NA	2	9
RF49b	LPS	19b	2	—	—	—	—
RF51	LPS	19a	2	J558	SP2.3,.4,.6	2	8
B13	LPS	23a	2	PC7183	SP2.1-.7	2	11
JV6	Alf-DNP	1c	1	J558	FL16.1,.2	4	9
VS1	KLH	1c	1	J606	Q52	3	8
A9	LPH-Ars	4a	2	PC7183	SP2.2 and FL16.1	2	13
JV9	Trf	4b	3	Sm7	NA	4	7
A12	LPH-Ars	8c	2	J558	FL16	2	8
A17	LPH-Ars	8d	2	J558	SP2.5,.7	3	9
JV10	Trf	8e	2	Q52	SP2.3,.4,.6	4	9
A23	LPH-Ars	19d	4	PC7183	SP2.5,.7,.8	2	11
A34	LPH-Ars	19d	1	J558	FL16.2	4	8
JV8	Trf	19d	3	PC7183	FL16.1′	3	12
VS2	KLH	19d	1	PC7183	FL16.1	1	10
A26	LPH-Ars	23a	1	PC7183	Q52	3	11
A15	LPH-Ars	24c	2	J558	SP2.5,.7,.8	4	8
A5	LPH-Ars	24b	1	J558	FL16.2	4	8

JV cell lines were produced by J. Van Snick; A and B lines by D. Nemazee. VS and RF cell lines were described previously, as were JV2 and JV6. Classification of the V region segments is as described in Results.

* $DxSO_4$, dextran sulfate; Alf-DNP, human apolactoferrin coupled with dinitrophenyl groups; Trf, human transferrin; LPH-Ars, arsonate conjugated to *Limulus polyphemus* hemocyanin.

‡ V_κ group number designations are according to Potter et al. (10); the small letters are arbitrary additions that distinguish different genes from the same family. For $V_\kappa 1$ and $V_\kappa 24$, the letters used correspond to the designations used by other authors for the same sequences when applicable (JV6, VS1, and A5).

§ J_κ genes are numbered 1–4 in the order 5′ → 3′ with respect to the constant region coding strand; the pseudogene (3 in other numbering systems) is skipped in this system.

‖ D_H designations are described in the text. The A9 V_H sequence has contiguous regions of homology to both SP2.2 and FL16.1. NA, not assignable to known gene.

¶ CDR3 length is defined as the number of amino acids between the last codon of FR3 (the invariant Arg residue) and the invariant Trp residue coded for by each of the J_H genes.

nucleic acid sequences are given in Figs. 3 and 4. The sequence of VS3 (8) is revised, and we report a partial sequence for the RF49 heavy chain.

To relate the RF V_H and V_κ sequences to other known sequences, we have classified each V region segment using the categories developed for V_H by Brodeur and Riblet (9) and for V_κ by Potter et al. (10) (Table I). V_H sequences are assigned to a group on the basis of ≥80% homology to prototype sequences from that family. By this criterion, the sequence of the V_H in JV9 (Fig. 2) defines

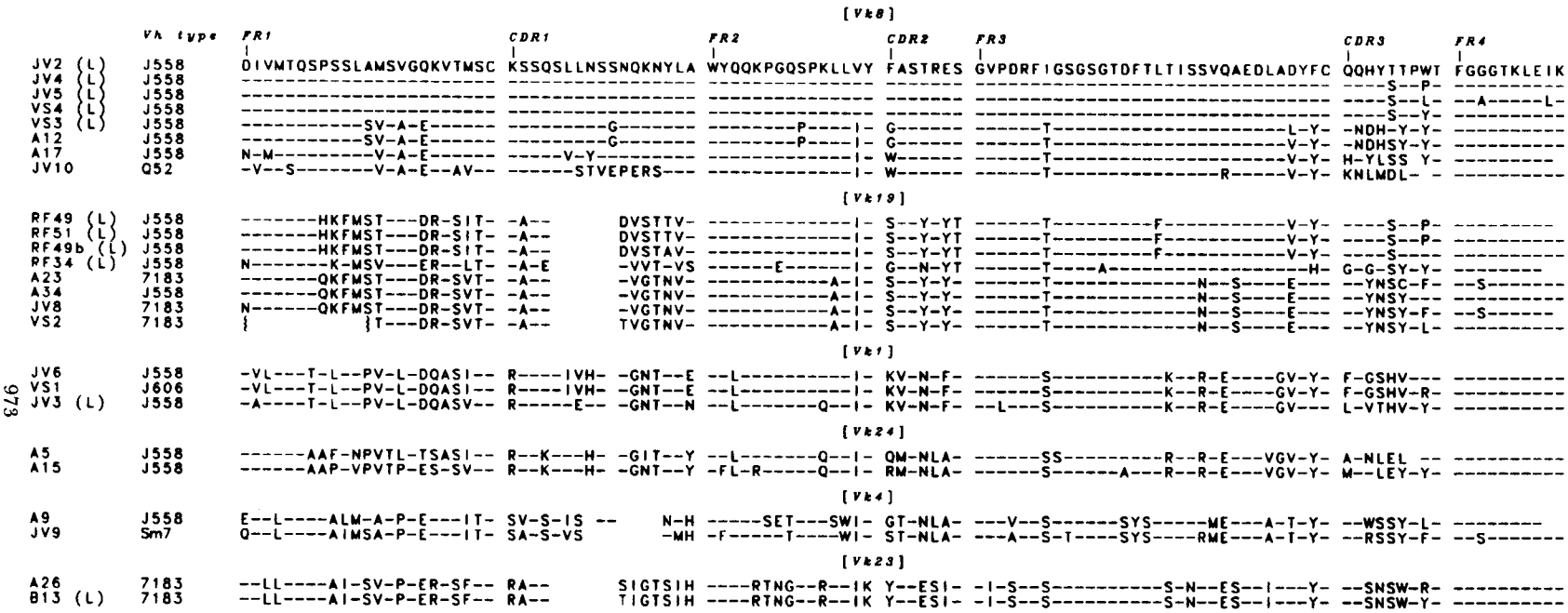

FIGURE 1. The light chain amino acid sequences of IgM anti-IgG1 hybridomas. Sequences were inferred from nucleotide sequences (Fig. 3) determined as described in Materials and Methods. Dashes indicate identity with the residue given in the first line. Cell names are at left, followed by the sequence on the same line. (*L*) after names denote LPS-derived hybridomas. RF49 expresses two productive V_κ mRNAs, both from the $V_\kappa 19$ family, which are designated RF49 and RF49b. The start of each framework (*FR*) or complementarity-determining region (*CDR*) (11) is given on the first line as FR or CDR followed by a number; these regions are set apart by a single space in the sequences. Blank spaces enclosed by curly brackets indicate regions of sequences that were not determined. Blank spaces are introduced elsewhere within sequences to maximize homology among sequences. Sequences belonging to the same V_κ group are given consecutively, followed by a space before the start of the next group, in this order: $V_\kappa 8$, $V_\kappa 19$, $V_\kappa 1$, $V_\kappa 24$, $V_\kappa 4$, and $V_\kappa 24$.

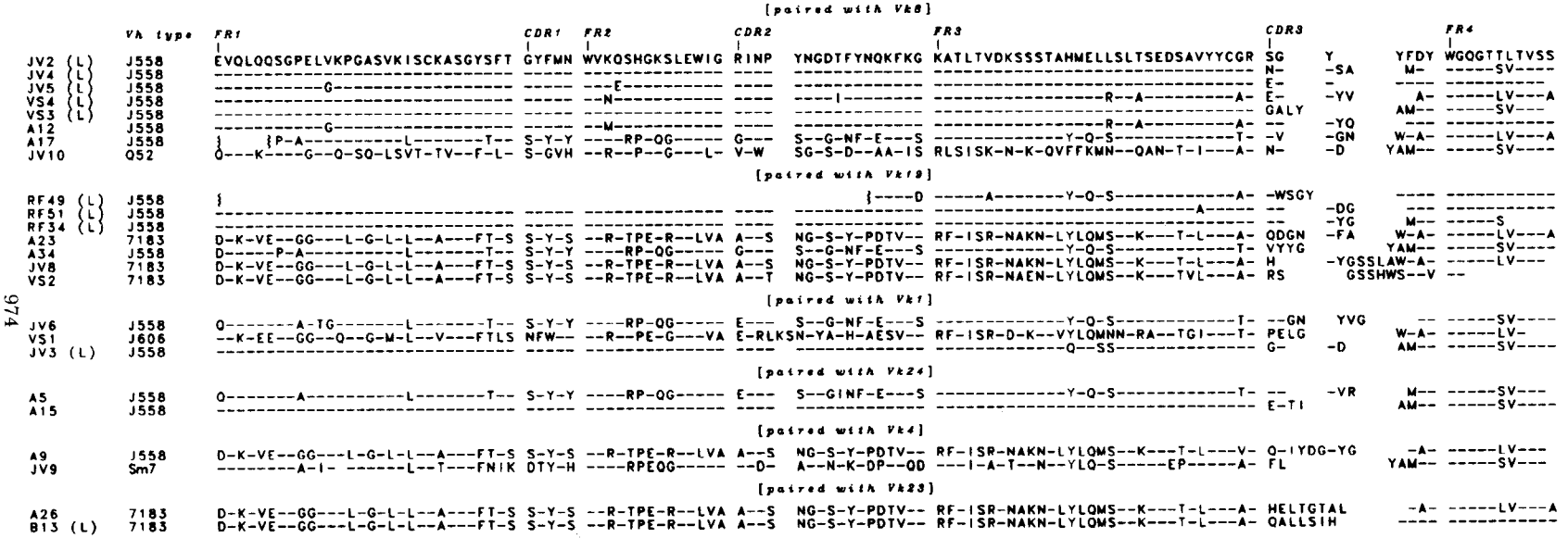

FIGURE 2. The heavy chain amino acid sequences of IgM anti-IgG1 hybridomas. The sequences are presented exactly as in Fig. 1. Thus, the heavy chain sequences are ordered according to the group of the light chain from the same cell line. See Table I for V_H family designation. The alignment of residues in CDR3 was done manually.

a new family, which we refer to as Sm7, after an anti-Sm hybridoma in which we first sequenced a V_H gene from this family (our unpublished data). Genes in this family do not have the invariant lysine residue at the 13th position, and are one residue shorter in the first framework region (FR1) (numbering and region designation as described in Kabat et al. [11]).

The V_κ groups described by Potter et al. (10) are based on amino acid sequence homology over the NH_3-terminal residues. A linking algorithm defined groups that had 12 or fewer differences over the first 41 residues. We have used this amino acid homology criterion to assign sequences to existing groups. Use of ≥80% nucleic acid homology would not have resulted in different assignments. Classification of J_H and J_κ is according to identity with the germline genes of BALB/c (12, 13). Classification of D_H is according to the best match with the known germline D_Hs of BALB/c (14); where ambiguity exists, all possibilities are listed in Table I.

Analysis of the nucleotide sequence data and Southern blot data (not shown) has allowed a tentative determination of whether related sequences belonging to the same V_κ group derive from the same V_κ germline gene. The data show that, for each V_κ group found among RFs (except V_κ23), multiple germline genes are used. The subclassification of the sequences according to presumptive germline genes is indicated in Table I. Sequences classified as coming from the same germline gene are either identical or have one (putatively somatic) difference at the nucleotide level. Genes classified as unique have 10–30 nucleotide differences from their closest relative; in many cases, two or more independent isolates of the same sequence confirms the distinction between it and its nearest relative. Detection of restriction enzyme map differences in Southern blotting analysis of representative hybridomas (not shown), supports the distinctions made by nucleic acid homology criteria. Among 25 RF V_κ sequences, we have observed four genes belonging to the V_κ8 homology group, four from the V_κ19 group, two from the V_κ1 group (15), two from the V_κ24 group (16), and two from the V_κ4 group.

The large size of the V_H J558 family (9), which may contain many nearly identical germline genes (17), prevents assignment of the J558 sequences on the basis of repeats. Unlike the V_κ sequences of RF, which contain many cases of repeats of the same sequence, there are 12 different V_H J558 family sequences among 15 RF examples (Fig. 1). On the other hand, the selection of genes from the 7183 family may be nonrandom, as only 2 different sequences are found in 6 examples; there are estimated to be 10–12 germline genes in this family. The possible implications of the patterns of V_H and V_κ gene expression in RFs will be discussed below.

Discussion

We have determined the complete V region sequences of 11 secondary protein immunization– and 5 LPS–derived RFs. These sequences, along with those already published, now provide a database large enough to permit a valid comparison of the two types of RFs. This comparison allows us to distinguish between models of the genetic control of RF production during the secondary immune response (see below). In addition, as discussed below, these data support

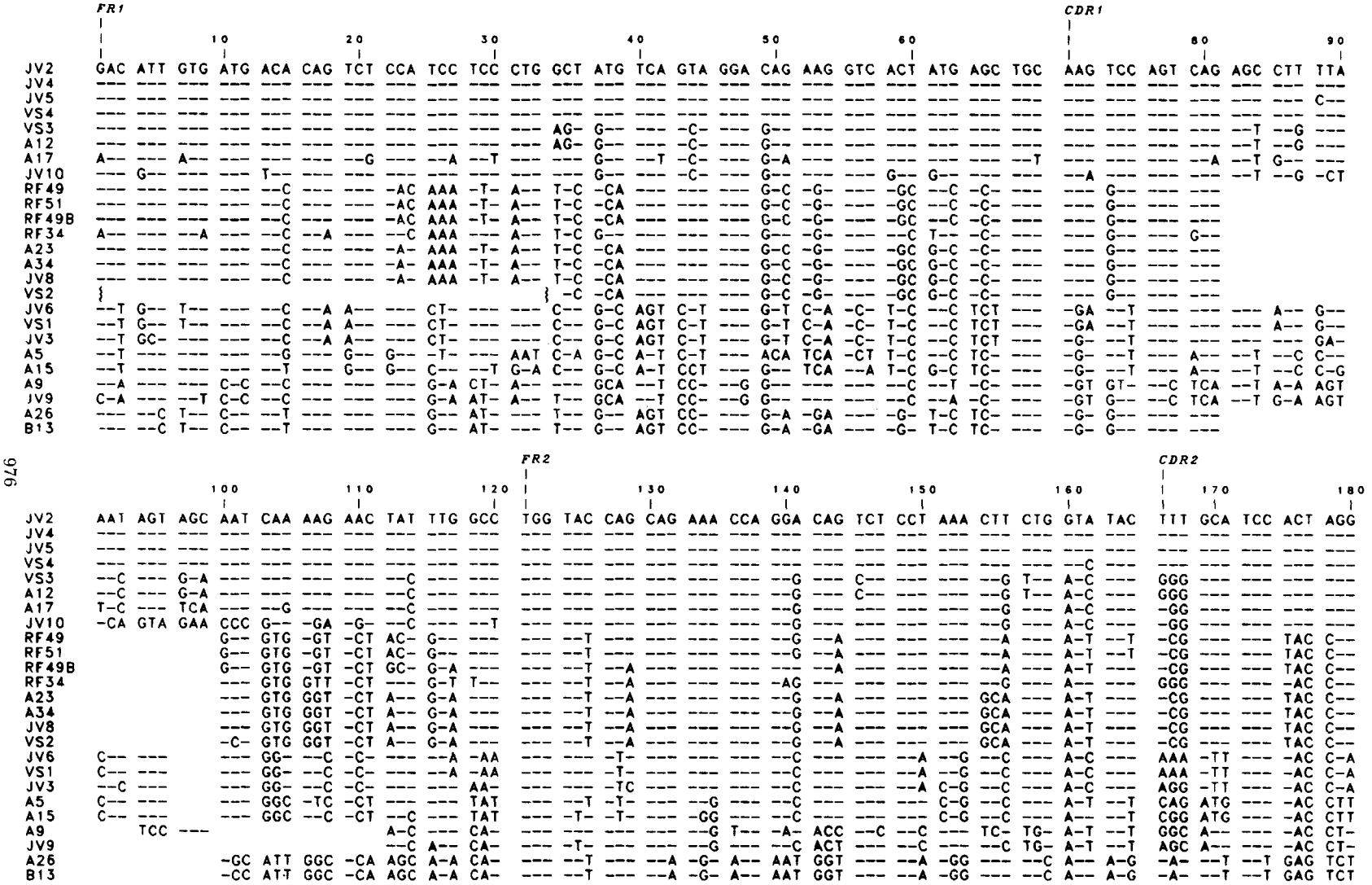

FIGURE 3. Nucleotide sequences of RF V_κ regions. Sequences are given in the same order as Fig. 1, and are displayed using the same conventions, except that V_H and V_κ designations and derivation are not indicated.

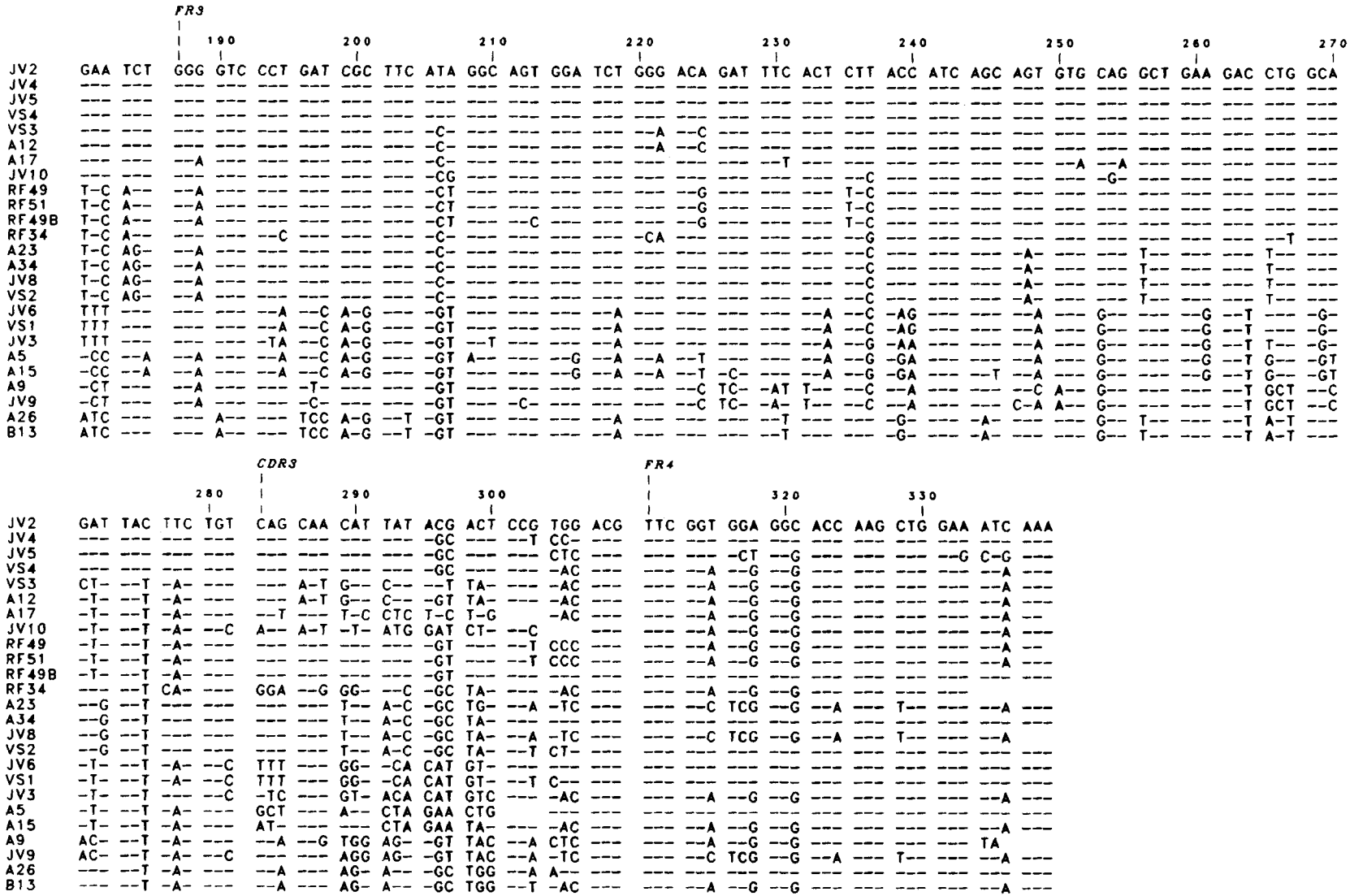

FIGURE 3—*Continued*

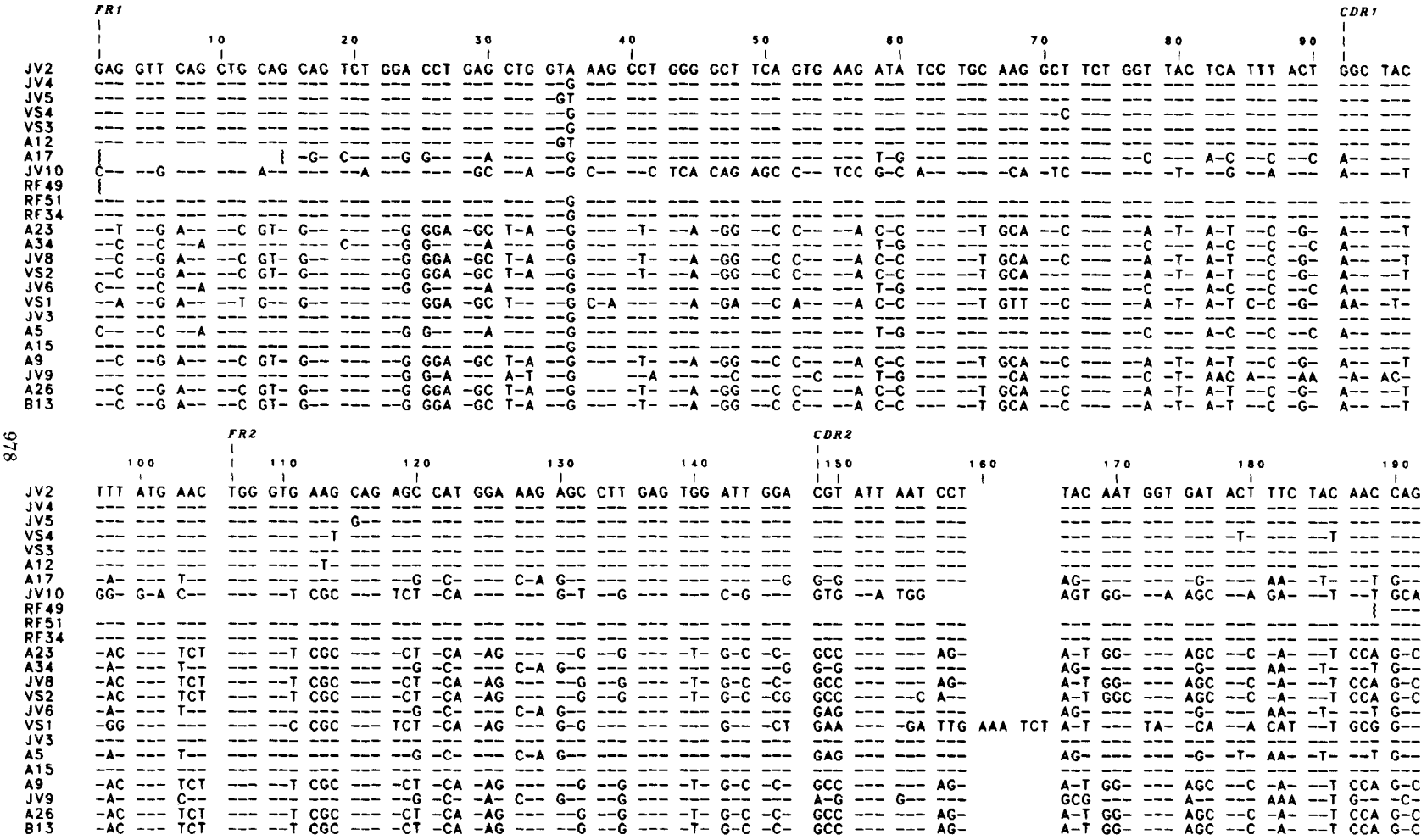

FIGURE 4. Nucleotide sequences of RF V_H regions. Sequences are given in the same order as Fig. 2, and are displayed using the same conventions, except that V_H and V_κ designations and derivation are not indicated.

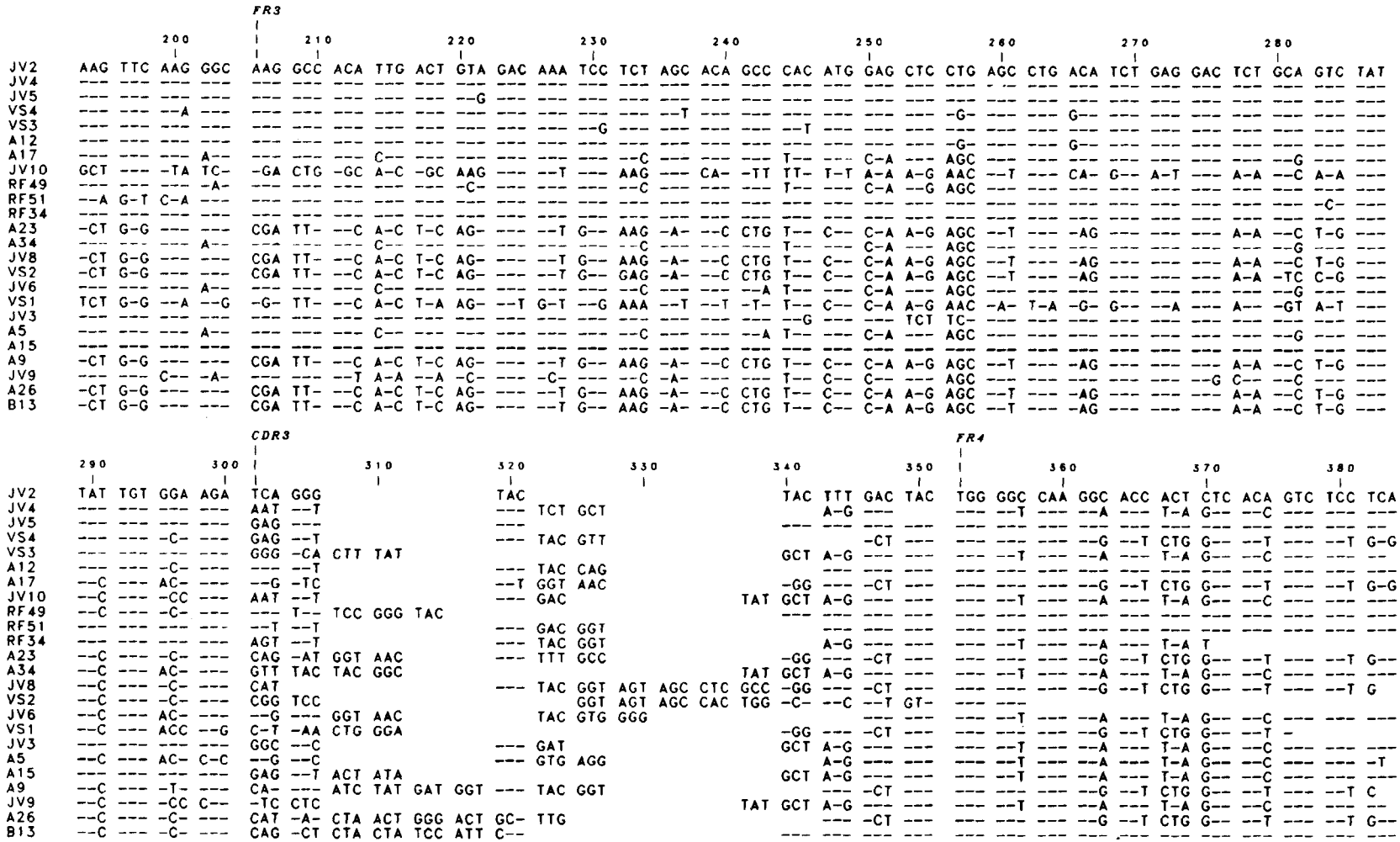

FIGURE4—*Continued*

our previous conclusions about V gene usage in RFs. Finally, we discuss the homologies among these V regions and how they define sites on RF molecules that may determine specificity.

The Pattern of V Region Expression in IgM Anti-IgG. Information about V region usage in RFs could provide insight into the reason for the high frequency of B cells that can make RF. Analysis of RF sequences for patterns of V region expression shows that a select subset of light chains is found, but there is remarkable diversity in the selection of the other variable region elements. This analysis consists of testing whether each V_H family or V_κ group is expressed more frequently in the RF population than would be expected randomly, using a binomial probability model. For each hybridoma, we assign the sequences to V_H groups and V_κ families, as described in Results and summarized in Table I. These classification systems reflect the V_H and V_κ gene organizations, which, briefly, are as follows: the V_H germline genes fall into ~10 sequence homology families, each containing from a few to perhaps hundreds of genes (9, 18). The V_HJ558 family appears to include at least 60% of the genes; the V_H7183, V_HJ606, and V_HQ52 families each contain at most 10% of the genes. The 100–200 V_κ genes are organized into ~30 groups, which each contain ~1–10 genes (10, 19). The level of expression for each group or family is not known precisely but is estimated from the fraction of germline genes that a given sequence group occupies. This estimate is reasonable because, in general, the expression of these genes in adults, as assayed in myelomas, LPS hybridomas, and serum, is concordant with their representation in the germline (20–22).

Statistical analysis of light and heavy chain representation shows that the selection of light chains, but not heavy chains, is significantly nonrandom among RFs. Each of three light chain types, V_κ1, V_κ8, and V_κ19 is found in our RF sample significantly more frequently than could be explained by random fluctuation ($p < 0.05$ for each). On the other hand, the representation of each of the V_H families is consistent with random selection of V_Hs from the available gene pool. The large number of V_HJ558-expressing RF hybridomas does not represent over-expression, because this single large family makes up ~60% of the unselected pool (21). Moreover, as inferred from sequence and Southern blot data (not shown) we have found at least 12 different members of the J558 family, suggesting that there is not selection for a particular J558 gene or subset of the J558 family. Genes from the V_H7183 family, which occur in 6 of 24 RFs, may be slightly overrepresented ($p = 0.03$) compared with the randomly expected value of ~10% (the approximate fraction of the germline gene pool that it occupies). We have so far observed only five V_H families, which together encompass 80–90% of V_H germline genes. The failure to observe the five other families may mean that not all V_Hs are suitable for generating anti-IgGs; but, because the gene families of the V_H types we have not observed are small, the results of the current survey could also be attributed to the sample size. The J_κs, D_Hs, and J_Hs include examples of all the known germline types, and there is no discernible pattern to their expression. The heavy chain complementarity-determining region 3 (CDR3) length varies from 7 to 12 residues. These latter elements do not seem to be conserved features of anti-IgG1 sequences.

Most important, highly divergent V_Hs from different families are found paired

with similar light chains for each of the dominant $V_\kappa 1$, $V_\kappa 8$, and $V_\kappa 19$ groups (compare VS1 and JV6, A17 and JV10, and A34 and JV8 V_H sequences, Fig. 2). Moreover, where more than one member of the same V_H sequence group (e.g., J558) is found paired to V_κs of the same family, these include multiple different members of that family (for example, compare JV3 and JV6, Fig. 2). The patterns of V region expression suggest a rule for the constitution of the anti-IgG1 specificity: the presence of any one of a subset of light chains ($V_\kappa 1$, $V_\kappa 8$, and $V_\kappa 19$), in combination with a variety of V_H, D_H, J_H, and J_κ seems sufficient to generate the capacity to bind IgG1. This rule can be used to make a rough estimate of the number of B cells that would rearrange the correct genes to make an RF antibody. This number can be calculated as the product of the fractions of V_H, D_H, J_H, V_κ, and J_κ regions that are suitable for the constitution of IgG binding. If one assumes, as suggested by the data, that a high fraction of V_H, D_H, J_H, and J_κ regions are suitable, then the percentage of B cells that have the RF specificity approaches the percentage of B cells expressing the correct V_κ genes. Using similar assumptions, we previously calculated this to be 5–17% of B cells. This compares favorably with the percentage of nonspecifically activated B cell hybridomas that make RF: 3–10%.

It is important to point out that this patten of V region expression is not typical of antibodies to other antigens. The most striking contrast comes from antibodies to small molecules (haptens) conjugated to protein carriers. Antibodies to haptens are often quite homogeneous, comprising restricted combinations of V region elements (V_H, D_H, J_H, V_L, and J_L) (17, 23). The extreme example of this is the dominant antibody of the A/J mouse immune response to *p*-azophenylarsonate (Ars), composed of a unique combination of V_H, D_H, J_H, V_κ, and J_κ segments (23). The likelihood of joining the required combination of segments in the same cell is low for such restricted responses. For the dominant anti-Ars response, it has been estimated that this event occurs in only 1 in 4×10^7 B cells (23).

RFs also seem different from antibodies to other protein antigens. Proteins can contain several different antigenic epitopes on their surface that are in some ways analogous to haptens. The immune responses to many protein epitopes have not displayed the same level of restriction as found for antihapten responses (24, 25). Typical restrictions that seem to govern the construction of antiprotein antibodies are illustrated by one set of hemagglutinin-specific antibodies (C4) frequently found in the primary response to influenza. In this case, the V_κ-J_κ combination is always the same, and the selection of V_H, though not always the same, is also clearly not random, in that three of five cases studied used a single germline gene (S. Clarke et al., manuscript in preparation). Thus, antibodies to some protein epitopes can be less homogeneous than antihapten antibodies but more homogeneous than IgM anti-IgG autoantibodies.

The Structures of RFs Derived from Nonspecific Activation Are Similar to Those Derived by Antigen Elicitation. The reason for making this comparison is to test the idea that the strong magnitude of the RF response during secondary protein immunization is related to the high number of B cells that can make that response, as suggested by the high frequency of RF-producing B cells activated by LPS. This idea implies that the two types are drawn from the same pool and predicts that the types of sequences found in LPS-derived and antigen-elicited RF anti-

bodies will be similar. This, however, need not be the case. An in vitro immunoassay was used to detect the LPS-derived hybridomas that made IgM anti-IgG, whereas in vivo B cell activation by antigen was the criterion for generating secondary protein immunization–derived RF. A second model states that possibly the antigen-antibody interactions measurable in such an in vitro assay would not be sufficient to cause B cell triggering in vivo. If this were the case, then one would expect that many of the structures found in LPS-derived RF would have no counterpart in the secondary protein immunization–elicited RFs, because these structures would not have sufficient affinity for IgG constant region to mediate B cell activation.

These sequence data are consistent with the first model. Each LPS-derived hybridoma in our sample has a counterpart, with similar V_H and V_κ sequences, among the group of secondary protein immunization–derived RF. 10 of 14 secondary protein immunization–derived RFs have sequences similar to those found in the LPS-derived RF sample. This suggests that most of the LPS-activated B cells that produce IgM anti-IgG could be precursors of the cells stimulated in vivo by IgG-containing immune complexes. Because it is known that a large fraction of LPS-activated B cells can produce RF, it is thus quite likely that the strong magnitude of the normal physiological RF response after secondary immunization with protein antigen is attributable to the presence of a large number of immune complex–activatible B cells.

For the foregoing analysis, it is only important to know whether the types of sequences found among LPS-derived RF are present among secondary protein-derived RF as well. A related question is whether the two groups are in fact the same in sequence composition. We have carried out χ^2 analyses to test this statistically. With the current data, we cannot rule out the null hypothesis that the LPS- and secondary, protein-derived RFs are the same in sequence composition for V_H ($p = 0.22$) or V_κ ($p = 0.42$) as a whole, nor can we rule out that the representation of any single V_H or V_κ type is the same between the two groups of hybridomas (all $p > 0.05$, Fisher's exact test).

Nonetheless, it is important to note that some V_Hs and some V_κs have been observed so far only in the secondary protein-derived hybridomas. Should, in a larger sample, these differences attain significance, it would be tempting to invoke selection by antigen as an explanation. It may be that RF stimulation by IgG-containing antigen-antibody complexes will result in the differential expansion of RF precursors; cells with higher-affinity receptors might be induced to divide more frequently. In this case, one might not expect the effects of selection for clones that bind antigen better to be very dramatic for RFs derived from animals immunized twice with protein, as the RF response in this context is analogous to a primary immune response. In fact, the data do not show strong selection effects; most (10 of 14) of the antigen-elicited RFs that we have sequenced do resemble the nonspecifically stimulated RFs. It will be interesting to see the effects of more protracted antigenic stimulation of RFs. Some diseased mouse strains carrying the *lpr* gene do make RF spontaneously, possibly as a result of continuous stimulation by immune complexes found in these mice. The population of RFs found in these mice may show stronger evidence of the

influence of selection by antigen. Determination of the primary structures of antibodies of this type should help to address this question.

Regions of Sequence that Are Conserved among RFs. Analysis of primary sequence data has often yielded useful information about the structure and function of antibodies. For example, Wu and Kabat (26) compared variable region sequences from a collection of unrelated Igs and noticed regions of primary sequence that were highly variable. These regions were inferred to be responsible for the unique specificity of individual antibodies. Since then, many others (17, 23, 24) have analyzed sequences of antibodies that bind the same antigen. Conversely, such antibodies of similar specificity proved to be identical at many residues that were hypervariable among randomly chosen antibodies. These shared residues were implicated in determining the shared antigen specificity; the importance of such residues has been shown directly in some cases (27, 28). Contact residues of Ig V regions are thus expected to be conserved among antibodies of the same specificity. The nonrandom pattern of V_κ representation and the fact that many hybridomas share very significant V_κ sequence homology, suggests that V_κ is playing an important role in determining the antigen specificity of RFs. Conversely, it is difficult to imagine, based on the extent of amino acid divergence among V_H sequences associated with V_κs of the 1, 8, and 19 groups, that the heavy chain is contributing to many, if any, contacts with antigen in these antibodies.

We therefore have focused on RF light chain amino acid sequences; previously we asked whether there are specific regions of sequence that are shared among many or all RF antibodies. To our surprise, we noticed that FR2 and FR3 of all sequences from V_κ1, V_κ8, and V_κ19 are highly similar compared with the same regions in V_κ sequences from other families (see Shlomchik et al. [8] for detailed analysis). The same analysis did not show any CDR or FR1 conservation between families. We therefore hypothesized that FR2 and FR3 may be involved in binding IgG1. This model is physically plausible, as these regions are on the surface of the V_κ domain. By this model, antibodies expressing V_κs from any one of these three families will bind IgG1 in a similar (though not necessarily identical) fashion, perhaps through an interaction between broad surfaces of the V_κ domain and an IgG1 constant domain (most likely CH3 [29]).

Further support for this model has emerged from three features of RF V region sequences that have become apparent from the present data. First is the fact that multiple germline genes from each of the V_κ groups can be expressed in RFs. Different genes in a V_κ family are often quite divergent in CDR but tend to resemble each other strongly in the FR regions (11). For example, compare the V_κ8 sequences of A17, JV10, VS3, and JV2 (Fig. 1). Thus, the pattern of FR2/3 conservation and CDR divergence noted in interfamily light chain comparisons is recapitulated in intrafamily sequence comparisons. Second, whereas previously we noted widely divergent V_Hs paired just with V_κ1 and V_κ19 chains, a prediction of the framework binding site model is that multiple types of V_H could also associate with V_κ8-type light chains in RFs. The newly determined sequences of JV10, in which a V_κ8 light chain is paired with a V_HQ52 heavy chain satisfies this prediction. Third is the finding of two V_κ24 light chains in antibodies A5 and A15. These two sequences (Fig. 1), although formally in the

$V_\kappa 24$ group, are quite similar to $V_\kappa 1$ sequences, particularly in FR2 and FR3 (as evident in Fig. 1). The levels of framework homologies are comparable to the homology between the two members of $V_\kappa 1$ itself. In addition, as found for $V_\kappa 1$, $V_\kappa 8$, and $V_\kappa 19$ RF sequences, the two observed members of the $V_\kappa 24$ family stem from two distinct germline genes and differ substantially in CDRs (Fig. 1).

Although V_κ is implicated as most important in supplying antigen contact residues, subtle features of the data do indicate a role for V_H in the determination of the specificity of these antibodies. Most striking is the example of JV4 and JV12, which express identical light chain sequences. The only structural differences between these antibodies are in the heavy chain, yet JV4 is specific for IgG1 while JV12 binds IgG2a of the a allotype but not IgG1 (JV12 sequence and binding data not shown). Although the antibodies use different members of the J558 family, it is notable that JV12 heavy chain has an unusually long CDR3, of 16 residues. Similar examples that bind IgG2a and not IgG1 occur with antibodies expressing $V_\kappa 1$ light chains and expressing $V_\kappa 19$ light chains (not shown). The effects of V_H on specificity need not, in principle, be mediated solely through direct antigen contact (27). Different V_H sequences could affect the folding or shape of light chains. Or, certain V_Hs might contain residues or CDR loop lengths that sterically prevent association of $V_\kappa 1$-, $V_\kappa 8$-, $V_\kappa 19$-, or $V_\kappa 24$-expressing antibodies with some types of IgG constant regions (e.g., JV12).

An alternative model that can account for V_κ restriction is that hybridomas expressing V_κ sequences from the same group bind in a similar way using the classical, CDR-encoded binding site. The lack of restriction on V_H and the use of multiple V_κ members from each family could be attributed to the relatively low affinity of RF antibodies, which might permit less stringent sequence requirements for these antibodies. In this model, each V_κ group would presumably define a different epitope on IgG1. While we do not favor this model for explaining the $V_\kappa 1$, $V_\kappa 8$, $V_\kappa 19$, and $V_\kappa 24$ sequences, we think it does account well for some of the infrequently observed sequence types. Hybridomas A26 and B13 express identical $V_\kappa 23$ sequences; use the same $V_H 7183$ gene, and have the same V_H CDR3 length. These could represent the type of antibodies that use CDR residues to bind a different epitope of the IgG1 constant region. A second pair of antibodies, JV9 and A9, use different genes belonging to the $V_\kappa 4$ light chain group and express V_Hs from two different families. Two other anti-IgG antibodies that use members of the $V_\kappa 4$ family use V_Hs from two additional families (our unpublished observations). This pattern of V_κ and V_H usage is reminiscent of that seen in the $V_\kappa 1$, $V_\kappa 8$, $V_\kappa 19$, and $V_\kappa 24$-expressing RFs. However, in FR2 and FR3, the $V_\kappa 4$ sequences are somewhat different from the other four V_κ groups. It is tempting to speculate that RFs expressing $V_\kappa 4$ genes represent another strategy for using FRs to bind IgG.

Regardless of the exact nature of the interaction between RF and IgG, the major conclusion of this study, that the two differently derived types of RF are structurally similar, has important implications for the etiology and function of RF. This result establishes a connection between the antibodies measured in the LPS response and those elicited by secondary immunization with protein antigen. The high frequency of these biologically relevant precursors could, in turn,

explain the very rapid and prolific nature of RF production accompanying a secondary immune response.

Thus, the capability to produce auto-IgM anti-IgG seems to be an inevitable result of the fact that murine RF antibodies are encoded by a relatively large set of V_κ germline genes in combination with a variety of V_H genes. The presence of these germline genes may indicate an evolutionary selection for and an adaptive value of the ability to produce large amounts of RF during a secondary immune response. Others have suggested roles for normal RF in enhancing complement fixation and immune complex clearance (2, 3), and have demonstrated its role in protection against *Trypanosoma lewisi* in rats (30). Further, in humans, as in the mice, there is evidence that RFs are encoded by a limited set of light chains in combination with a variety of heavy chains (31). Here, too, a similar strategy for providing a high frequency of B cells that express what might be a critical specificity may be used.

Summary

We have obtained the complete variable region mRNA sequences of 11 LPS-derived and 14 secondary immunization–derived monoclonal IgM anti-IgG antibodies (rheumatoid factors, RFs). A comparative analysis of these sequences showed that monoclonal RFs derived after polyclonal activation are structurally very similar to RFs derived after secondary protein immunization. This study was undertaken to evaluate the potential relationship between two previously described phenomena: (*a*) during a secondary response to a protein antigen, RF is produced in quantities that equal or exceed the immunogen-specific antibody; and (*b*) the frequency of B cells that make RF after polyclonal activation is quite high; 3–10%. It has been unclear whether LPS-stimulated cells that produce IgM anti-IgG that is detected by an in vitro assay are related to the cells that produce RF after in vivo stimulation. The similarity of the antigen receptors found in the two types of RF, however, suggests that most or all of the RF-producing B cells detected after LPS stimulation would also be stimulated during the secondary immune response. Thus, the presence of relatively large number of B cells that can make RF after nonspecific stimulation provides an explanation for the magnitude of RF production accompanying the secondary immune response.

We thank Anita Cywinski for superb technical work, Sam Litwin for discussion of statistical analysis, and Ann Haberman for criticism of the manuscript.

Received for publication 9 October 1986 and in revised form 4 December 1986.

References

1. 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.
2. 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.
3. 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.

4. Rose, H., C. Ragan, E. Pearie, and M. Lipman. 1948. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. *Proc. Soc. Exp. Biol. Med.* 68:1.
5. Nemazee, D. 1985. Immune complexes can trigger specific, T cell–dependent, autoanti-IgG antibody production in mice. *J. Exp. Med.* 161:242.
6. 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.
7. 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.
8. Shlomchik, M., D. Nemazee, V. Sato, J. Van Snick, D. Carson, and M. Weigert. 1986. 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. *J. Exp. Med.* 164:407.
9. 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.
10. 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.
11. 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.
12. Max, E., J. Seidman, and P. Leder. 1979. Sequences of five potential recombination sites encoded close to an immunoglobulin κ constant region gene. *Proc. Natl. Acad. Sci. USA.* 76:3450.
13. Gough, N., and O. Bernard. 1981. Sequences of the joining region genes for immunoglobulin heavy chains and their role in the generation of antibody diversity. *Proc. Natl. Acad. Sci. USA.* 78:509.
14. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *J. Exp. Med.* 155:201.
15. Hum, W.-T., C. Lazure, A. Lavigueur, S. MacLean, and D. Gibson. 1984. Structure and genetic control of V_κ-1 light chains. *Ann. Immunol. (Paris).* 135C:163.
16. Joho, R., H. Gershenfeld, and I. Weissman. 1984. Evolution of a multigene family of V_κ germline genes. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:185.
17. Bothwell, A. L. M., M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, and D. Baltimore. 1981. Heavy chain variable region contribution to the NP^b family of antibodies: Somatic mutations evident in a γ_{2a} variable region. *Cell.* 24:625.
18. Goverman, J., T. Hunkapiller, and L. Hood. 1986. A speculative view of the multicompetent nature of T cell antigen recognition. *Cell.* 45:475.
19. Weigert, M., and R. Riblet. 1976. Genetic control of antibody variable regions. *Cold Spring Harbor Symp. Quant. Biol.* 41:837.
20. Gibson, D. M. 1984. Evidence for 65 electrophoretically distinct groups of light chains in BALB/c and NZB myelomas. *Mol. Immunol.* 21:421.
21. Dildrop, R., U. Krawinkel, E. Winter, and K. Rajewsky. 1985. V_H gene expression in murine lipopolysaccharide blasts distributes over the nine known V_H groups and may be random. *Eur. J. Immunol.* 15:1154.

22. 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.
23. 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.
24. Caton, A. J., G. G. Brownlee, L. M. Staudt, and W. Gerhard. 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.
25. 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.
26. Wu, T. T., and E. Kabat. 1970. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* 132:211.
27. Davies, D. R., and H. Metzger. 1983. Structural basis of antibody function. *Ann. Rev. Immunol.* 1:87.
28. Sharon, J., M. L. Gefter, T. Manser, and M. Ptashne. 1986. Site-directed mutagenesis of an invariant amino acid residue at the variable-diversity segments junction of an antibody. *Proc. Natl. Acad. Sci. USA.* 83:2628.
29. 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.
30. 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.
31. 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.