Multiple Differences between the Nucleic Acid Sequences of the IgG2aa and IgG2ab Alleles of the Mouse

Peter H. Schreier, Alfred L. M. Bothwell, Benno Mueller-Hill, and David Baltimore

PNAS 1981;78;4495-4499
doi:10.1073/pnas.78.7.4495

Proc. Natl. Acad. Sci. USA
Vol. 78, No. 7, pp. 4495–4499, July 1981
Genetics

Multiple differences between the nucleic acid sequences of the *IgG2a*^{a} and *IgG2a*^{b} alleles of the mouse

(evolution/polymorphism/immunoglobulin heavy chain genes)

PETER H. SCHREIER*, ALFRED L. M. BOTHWELL†, BENNO MUELLER-HILL*, AND DAVID BALTIMORE†

*Institute for Genetics, University of Cologne, Weyertal 121, D-5000 Köln, West Germany; and †Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Contributed by David Baltimore, April 3, 1981

ABSTRACT **To compare the structure of *IgG2a* alleles we have determined the complete DNA sequence of the constant region, coding sequence, and 3′ untranslated region of a cDNA clone, pABγ2a-1, which was derived from the C57BL/6 mouse strain (*b* allotype). This sequence was compared with the corresponding *IgG2a* DNA sequence of BALB/c origin (*a* allotype). The DNA sequences showed 10% differences, and the deduced protein sequences differed by about 15%. These differences were not evenly distributed: most differences were in the hinge region, the C_H3 domain and the 3′ untranslated region. It is evident that many alterations in the *IgG2a* alleles have occurred since the *a* and *b* haplotypes were separated—some of these changes were point mutations but some appear to have resulted from gene conversion of the *IgG2a*^{b} allele by the *IgG2b*^{b} allele.**

The immunoglobulin heavy chain loci of the mice are inherited as *Igh* haplotypes—tightly linked genes that rarely reassort by recombination and that encode the constant (C) region segments (1). Many *Igh* haplotypes are found among the various inbred strains of mice; they are differentiated by the presence of allotypic markers, distinct antigenic determinants present on homologous proteins that result from genetic polymorphism (2). BALB/c mice have the *Igh*^{a} haplotype; C57BL/6 mice have the *Igh*^{b} haplotype. The number of constant region genes in a haplotype is generally thought to be at least 8—encoding μ, δ, γ3, γ1, γ2b, γ2a, ε, and α in that order (1)—but the proof that this number is correct or that the number or sequence is invariant among inbred strains is lacking. Only in BALB/c mice, in which Shimizu *et al.* (3) have linked by molecular cloning methods a segment from the gene encoding γ3 to that encoding α, is the sequence firmly established.

It is thought that most inbred strains have a common set of constant regions because serologic reagents can be used to establish homologies across strains. γ2a and γ2b proteins have been identified in the sera of both BALB/c and C57BL/6 mice. The γ2a proteins of these strains, however, are known to differ extensively: Oi and Herzenberg (4) have raised six monoclonal antibodies that differentiate the two proteins and Seman *et al.* (5) found 3 amino acid differences among 30 positions examined.

To examine in more detail the differences between these two apparently allelic *IgG2a* genes, we have determined the DNA sequence of the *IgG2a*^{b} allele and compared it to the published data on the *IgG2a*^{a} allele. Surprisingly, about 10% of the nucleotide positions in the sequences of these allelic genes differed. The predicted γ2a^b amino acid sequence, however, was almost identical to the amino acid sequence for the γ2a^b protein determined independently (6).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

T4 ligase was isolated (7) from an *Escherichia coli* strain carrying λ T4 Lig (989) (8) kindly provided by N. Murray. All of the restriction endonucleases were obtained from New England BioLabs. The Klenow fragment of *E. coli* polymerase I was obtained from Boehringer Mannheim. *E. coli* K-12 strain 79-02 which was provided by B. Gronenborn is a tra-D36 derivative of strain 79-01 which was derived from *E. coli* K-12 strain C600 and is (r^-k m^-k, Thr, Leu, Str^R, lacZ,Y, Pro, F′ lac I^qZ M15 Pro^+). The M13mp2Bam phage was provided by R. Cortese and is a derivative of a phage, constructed by Rothstein *et al.* (9), that forms blue plaques on strain 79-02 growing on plates containing 5-bromo-4-chloro-indolyl-β-D-galactoside and isopropyl thiogalactoside. This phage additionally contains no *Sma* I site but has pA-A-T-T-C-C-C-C-G-G-A-T-C-C-G-G-G-G inserted in the *Eco*RI site of M13mp2.

Recombinant DNA Plasmids. The plasmid pABγ2a-1 was derived from the S43 hybridoma (10) which was a fusion product of the BALB/c myeloma cell line X63-Ag8 and a spleen cell from a hyperimmunized C57BL/6 mouse. This plasmid contained the entire coding sequence for the C57BL/6 γ2a heavy chain. The details of the preparation and the structure of the plasmid will be published elsewhere (11). For subcloning in M13mp2Bam, pABγ2a-1 DNA was digested with *Sau*3A and the fragments either were ligated directly into the vector or, prior to ligation, were isolated on a 2% low-melting agarose gel. Electrophoresis was carried out in 40 mM Tris acetate/20 mM sodium acetate/1 mM EDTA, pH 8.0. DNA was extracted from the agarose by melting the gel slice at 70°C until no residual gel was visible and subjecting it to high speed centrifugation after the solution was cooled to 0°C. The supernatant was phenol-extracted twice, the residual phenol was extracted with ether, and the DNA was ethanol-precipitated and dried. Recombinant phage were screened and templates were isolated as described (12). For DNA sequence determination, current methods were used (13–15) and are discussed in *Results and Discussion*. Subcloned fragments in M13mp2Bam were analyzed by the chain termination procedure (15, 16) as described (12). Restriction fragments of pSP14 (17), mp2962 (12), and M13.2a-16 were used as primers. Sequence data were compiled with computer programs of Staden (18) which were modified and made compatible to a ZILOG Z80 computer by K. Stueber.

RESULTS AND DISCUSSION

Nucleotide Sequence of the C57BL/6 γ2a Heavy Chain Constant (C_H) Region. The nucleotide sequence of the *IgG2a*^{b}

Abbreviations: C region, constant region; C_H, C region of heavy chain; bp, base pair(s); V region, variable region.

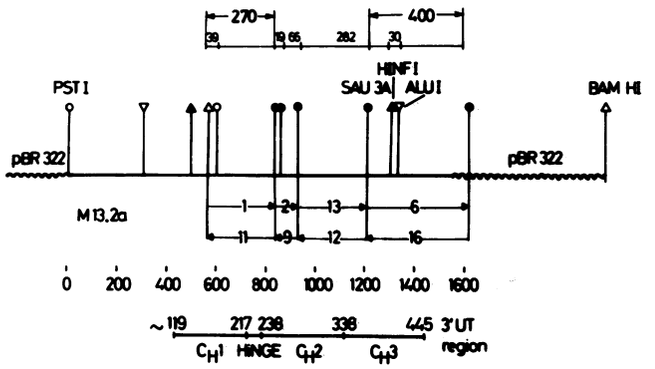

FIG. 1. Sites of restriction enzyme cleavage in the cDNA insert in pABγ2a-1. From top to bottom: size of restriction fragments; restriction map; M13 subclones; insert base pairs; and protein residues and domains. Only restriction sites directly involved in the sequencing strategy are shown. The nucleotides are numbered from 5′ to 3′ on the coding strand, beginning with the residue corresponding to the first nucleotide of the leader sequence. The C_H1 domain begins with amino acid 119. The wavy lines on each side of the insert represent the synthetic dG-dC residues and the pBR322 vector DNA. The subclones of pABγ2a-1 in M13mp2BAM are designated M13.2a-1, -2, -6, -9, -11, -12 and -16. The arrows indicate the orientation of the insert in the vector corresponding to sequence obtained either from 5′ to 3′ in the coding strand or from 5′ to 3′ on the noncoding strand. The first 150 bp in C_H1 of the coding strand sequence were determined by using DNA fragments end-labeled at the *Alu* I and *Hinf* I sites at approximately 300 and 500 bp within the cloned insert. The sequence of the noncoding strand was determined by end-labeling at the *Pst* I site (position 600) and at an *Hph* I site located at position 500 in the cloned insert. For subcloning in the M13 derivative a *Bam*HI fragment from plasmid pABγ2a-1, of about 2000 bp, was isolated from a low-melting agarose gel and further digested with *Sau*3A. The resulting fragments were ligated to M13mp2Bam and hybrid phage was screened on plates containing 5-bromo-4-chloro-indolyl-β-D-galactoside and isopropyl-thiogalactoside (19). Of 20 hybrid clones analyzed, only 4 contained DNA from the cDNA insert—namely, the small *Sau*3A fragments containing subfragments of 19 and 66 bp in both orientations (M13.2a-2 and -9). The 16 others contained small *Sau*3A fragments from pBR322 DNA also present in the 2000-bp insert. To clone the remaining fragments, a *Sau*3A fragment of 400 bp and two *Sau*3A fragments of 270 and 282 bp were isolated and cloned in two separate experiments. Two clones (M13.2a-6 and -16) from the former and four clones (M13.2a-1, -11, -12 and -13) from the latter experiment produced hybrid phage containing the remaining *Sau*3A fragments of the γ2a cDNA in both orientations. The DNA sequence was obtained by the chain termination technique of Sanger and coworkers (15, 16) adapted to the M13 system as described (12). To obtain the first nucleotides of each insert, a primer of 96 nucleotides which hybridizes adjacent to the insert was pretreated with exonuclease III. The sequence generated spanned the 3′ end of the primer, the cloning site, and the beginning of the insert (12). Most of the other sequencing experiments were done with a single primer (20) of 30 bp from plasmid pSP14 (17) which also hybridizes adjacent to the insert. The *Sau*3A recognition sequence, G-A-T-C, indicated the termini of the inserted DNA. An exception was M13.2a-1 which had a *Bam*HI site at its 5′ end. This enabled us to align the fragments without any gaps between fragments. The unambiguous order of the fragments was rendered possible by the protein sequence (21) and the cDNA sequence (22) of the *IgG2a*^{a} allele. We were not able to obtain clear sequence data from M13.2a-16 with the primer of 27 bp from pSP14 beyond the poly(G) and poly(A) tails. This suggested that there was intramolecular base pairing of the single-stranded DNA in this region, which prevented electrophoretic mobility according to size (23). Some DNA sequence of the noncoding strand from M13.2a-16 was determined by using a 30-bp primer derived from supercoiled DNA of M13.2a-16 by *Hinf* I and *Alu* I digestion.

C_H region gene was determined by using plasmid pABγ2a-1 (Fig. 1) which contains an insert of about 1600 base pairs (bp) derived from the C57BL/6 *IgG2a* contribution to the hybridoma S43 (11). The sequence of the 600-bp *Pst* I fragment encoding the variable (V) region, D and J regions, and the first approximately 150 bp of the C_H1 domain of the S43 γ2a protein was determined on both strands by using Maxam and Gilbert procedures (13). The C region DNA sequence from that determination is presented here. The sequence of the rest of the IgG2a C region was determined by using subcloned restriction fragments in a derivative of the single-stranded DNA phage, M13 (Fig. 1).

Comparison of *IgG2a*^{b} (C57BL/6) and *IgG2a*^{a} (BALB/c) Sequences. The sequence of the γ2a C region DNA from the C57BL/6 genome is presented in Fig. 2. It contains 1108 bases representing the C_H1, hinge, C_H2, and C_H3 domains and the 3′ untranslated region of the *IgG2a*^{b} gene. (Because of insertions and deletions, the sequence is numbered, along with the *IgG2a*^{a} and *IgG2b*^{a} sequences, as having 1114 bases in Fig. 2.) Comparison of this sequence and that of the *IgG2a*^{a} gene (22) shows a total of 111 differences plus 15 additional nucleotides in the hinge region. In C_H3, the CTT of codon 430 is absent from the *IgG2a*^{a} gene. This gene has an additional GTG codon inserted after codon 424, thus conserving the reading frame and the length of the domain. The amino acid sequence of these proteins in the C_H3 regions also required the same insertion and deletion to maintain homology (6).

The amino acid sequence deduced from the nucleotide sequence is presented in Fig. 3. The deduced sequence of the S43 protein differs at only two positions (387 and 449) from the C_H2 and C_H3 sequences of the $IgG2a^b$ protein from CBPC 101 cells (6). Assuming no sequencing errors, the small differences could be explained by somatic mutations in the hybridoma or myeloma lines or possibly genetic polymorphisms. The homology between the deduced sequence of the S43 protein and the determined sequence of the CBPC 101 protein (6) indicates that we are actually dealing with a germ-line sequence and not with extensive somatic mutations of the *IgG2a* gene, although somatic mutations have been shown to occur in the V_H gene which is linked to this C gene (11).

The differences between the protein sequences encoded by the *IgG2a*^{b} and *IgG2a*^{a} genes (24) are quantitated in Table 1 by domain. The two alleles differ only slightly more than do the two neighboring genes of the same haplotype (*IgG2a*^{a} and *IgG2b*^{a}). In fact, in the hinge region the alleles differ in length by six amino acids but the neighboring genes differ by only one amino acid. Compared by domain, the three sequences have similar distributions of nonhomology. The most conserved regions of sequence among the three genes are in C_H2 and C_H1; C_H3 and the hinge region vary the most.

Replacement and Silent Differences. Of the 111 differences between the *IgG2a*^{a} and *IgG2a*^{b} alleles, 18 (16.2%) are silent differences and the rest lead to amino acid changes (they are replacement changes). Because the random expectation of silent differences is 24.4%, it is evident that there has not been extensive selection against replacements; in fact, the data suggest that there may have been selection for replacement. It should be emphasized that the alleles in question derive from one species of animal and thus the arguments for selection of replacements during speciation (24, 25) are not obviously relevant. Whatever the explanation for the high level of replacement changes, it indicates that the sequence alterations are of recent origin because widely separated genes generally accumulate many silent differences (25).

Gene Conversion or Unequal Crossing-Over. One of the most striking results from comparing the *a* and *b* alleles of the *IgG2a* gene with the *IgG2b*^{a} allele (Figs. 2 and 3) is that in three regions where *IgG2a*^{a} differs from *IgG2a*^{b}, the differences are present in the *IgG2b*^{a} sequence. Between codons 123 and 135 there are five nucleotide positions where the *IgG2a*^{a} and

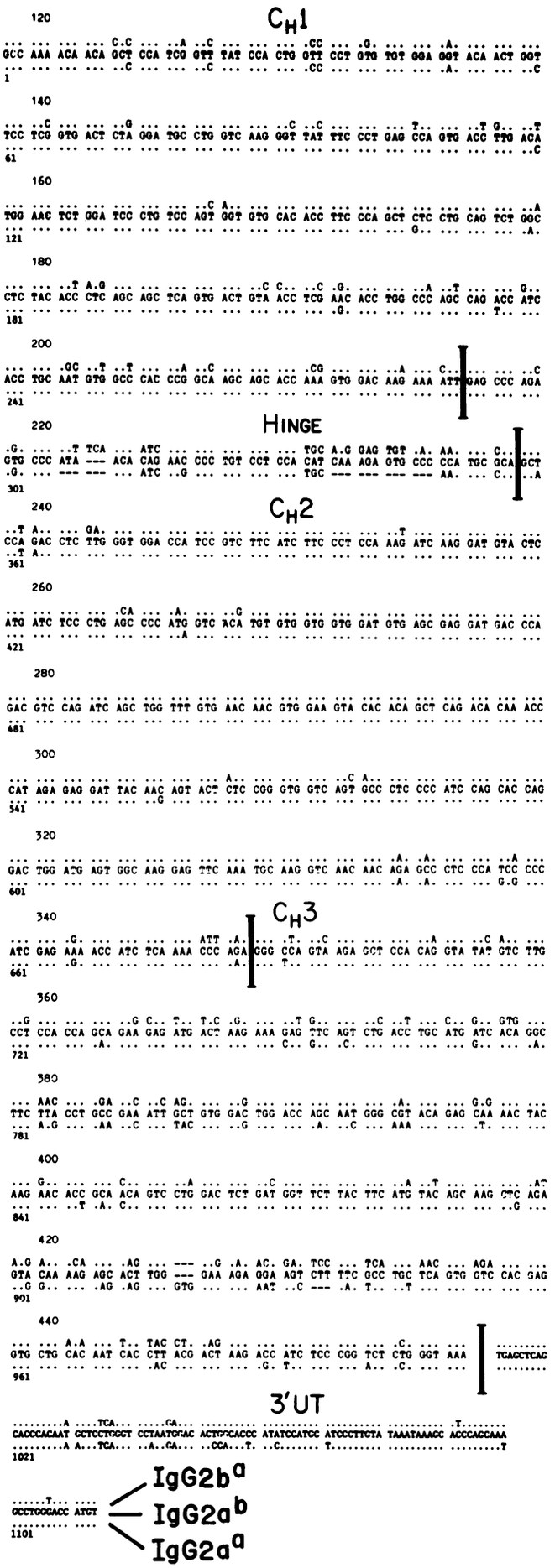

FIG. 2. Comparison of nucleotide sequences of the *IgG2a*^{a}, *IgG2a*^{b}, and *IgG2b*^{a} C regions. The sequence is numbered from 5′ to 3′ on the coding strand. The first codon shown corresponds to residue 119 of the H chain. Only the nucleotide sequence of *IgG2a*^{b} is fully printed. Identical nucleotide sequences shown above and below this sequence are indicated by dots. Differences are indicated. Gaps used to align sequences are shown by dashes. The sequence of *IgG2a*^{a} and *IgG2b*^{a} were determined by Sikarov *et al.* (22) and Bothwell *et al.* (11), respectively.

IgG2a^b alleles differ; all of the five nucleotides specific to the *IgG2a*^{a} sequence are present in the *IgG2b*^{a} sequence at these positions. In the area of the hinge region, excluding the gaps, from codons 219 to 240 there are 14 differences between *IgG2a*^{a} and *IgG2a*^{b} of which 12 are positions where *IgG2a*^{a} and *IgG2b*^{a} are identical. Finally, in the 3′ external region, from nucleotide positions 1030 to 1047 there are 8 differences between *IgG2a*^{a} and *IgG2a*^{b} of which 6 are identities between *IgG2a*^{a} and *IgG2b*^{a}. Aside from these three localized regions, over the segment from codon 333 to codon 453 (the whole C_H3 domain plus a bit of C_H2), about one-third of the 51 differences between *IgG2a*^{a} and *IgG2a*^{b} are identities between *IgG2a*^{a} and *IgG2b*^{a}. By contrast, from codon 138 to codon 304, but excluding the hinge area, there are only 8 scattered polymorphisms of *IgG2a* and only one is an identity of *IgG2b*^{a} and *IgG2a*^{a}.

There are two possible interpretations of these segmental homologies between nonallelic but neighboring genes: the genes could have participated in gene conversion events, or they could have duplicated one from another. The duplication hypothesis is equivalent to postulating unequal crossing-over events between sister chromatids (26).

Gene conversion is a process in which gene A interacts with gene B in such a way that a region of the sequence of gene B is converted to be identical in sequence to gene A (a model is presented in ref. 27). Such events have been classically described for allelic genes in fungi (28). Gene conversion events differ from unequal crossing-over events in that unequal crossing-over increases or decreases the number of genes in a family of related genes whereas gene conversion events do not change the number or the order of genes (just a local region of base sequence is altered).

Without knowing precisely the number of *IgG2a*-like genes in the C57BL/6 genome or having any basis other than the serologic data to say that the *IgG2a* alleles are truly allelic, it is difficult to choose with certainty between unequal crossing-over and gene conversion as explanations for the homologies between nonallelic genes. A number of arguments, however, support gene conversion as the explanation. First, unequal crossing-over events would increase or decrease the number of IgG-like C regions; this could be deleterious to the functioning of the immune system and, if so, then gene conversion is a more likely origin for the *cis* gene homologies. Second, Slightom *et al.* (27) have observed segmental homologies between neighboring β-globin-like genes and have argued in favor of gene conversion as a mechanism to explain their data—by the same arguments, our observation of segments of high homology between *IgG2a*^{a} and *IgG2b*^{a} separated by regions of low homology favor a gene conversion model. Miyata *et al.* (29) have observed segmental homologies between the *IgG1* and *IgG2b* genes of BALB/c mice and have argued for unequal crossing-over events as the explanation for the homologies; their data could as well result from gene conversion events and such an explanation would avoid the complex recombinational events they postulate. As emphasized elsewhere (30), gene conversion events also could explain the evidence that led Kabat *et al.* to the "minigene hypothesis" (31, 32) and could explain the shared immunologic specificities of *H-2D* and *H-2K* gene products (33, 34).

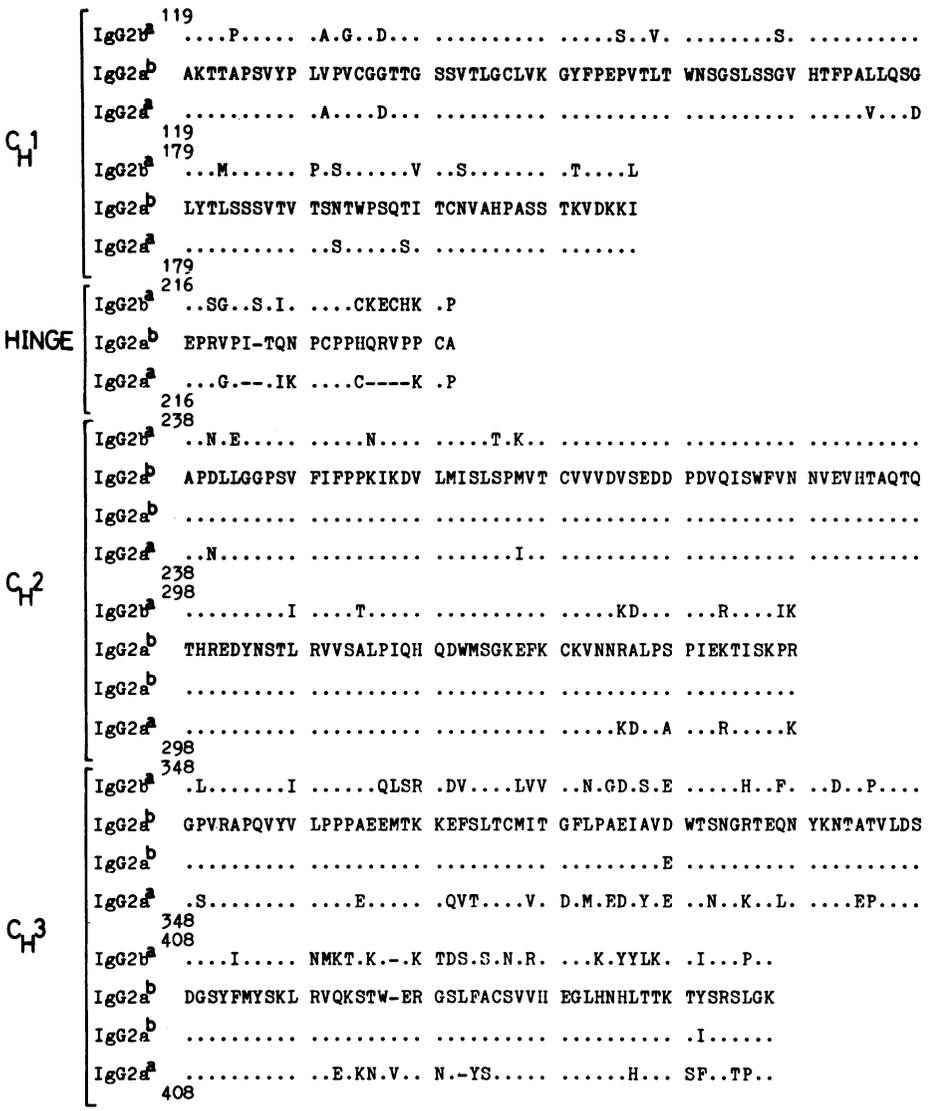

FIG. 3. Comparison of predicted amino acid sequences of the IgG2a^a, IgG2a^b, and IgG2b^a proteins. Only the amino acid sequence of IgG2a^b is fully printed. Identical amino acid sequences shown above and below this sequence are indicated by dots. Differences are indicated. Dashes are inserted to maximize homology. The numbering starts at residue 119 in C_H1 and proceeds from the amino to the carboxyl terminus.

Aside from the occurrence of positions of intergene homologies, there are many positions where *IgG2a^a* and *IgG2a^b* differ. These could have resulted from single base changes in one or the other allele, but then the two alleles must have evolved separately for a long time (6, 25). More likely, gene conversion or other recombination events with as yet unidentified genes are responsible for the divergences. The level of replacement differences between these two alleles is what one might expect for alleles in the major histocompatibility complex (35, 36)—whether or not the selective forces at work have any relationship is unclear.

There may be selective maintenance of sequence in those areas of the gene where few alterations are found. It is relevant that those amino acids mentioned in a study on the possible binding site of C1q in C_H2 (37) were conserved in both of the allotypes. Furthermore, none of the residues mentioned for binding of carbohydrates or the maintenance of specific structural features were altered.

Assignment of Allotypic Determinants. Serological experiments (4) on proteolytic fragments of the IgG2a^b molecule have shown that four allotypic determinants are located on the following parts of the molecule: one in or near the hinge, one in the C_H2 domain, and two in the C_H3 domain. By virtue of the large number of differences found, we were unable to assign single amino acids to allotypic determinants. However, we can narrow down the possible amino acids or combinations of amino

Table 1. Comparison of homologies between IgG2a^a, IgG2a^b, and IgG2b^a C_H domains

	C_H1		Hinge		C_H2		C_H3		Entire C_H	
Example	Homology, %	Gaps, no.	Homology, %	Gaps, no.	Homology, %	Gaps, no.	Homology, %	Gaps, no.	Homology, %	Gaps, no.
IgG2a^a vs. IgG2a^b	94	0/0	71	2/5	94	0/0	72	0/0	85	2/5
IgG2b^a vs. IgG2a^a	86	0/0	72	2/6	91	0/0	63	0/0	80	2/6
IgG2b^a vs. IgG2a^b	86	0/0	45	1/1	90	0/0	61	0/0	77	1/1

Percentage of homology = 100 × [amino acid identities/(amino acids compared + number of gaps)]. Gaps in the alignment are counted as single differences. Under "gaps" are shown the number of gaps required to maximize the sequence homology and the total number of amino acids included in the gaps.

acids responsible for the determinants to be detected by alloantisera. The five different amino acids in the hinge or more likely the five additional residues in the hinge are likely to be involved in forming the first determinant. Only the amino acids in positions 240, 265, 333, 334, 337, 341, and 347 and probably in 349 are likely to be involved in the second determinant. With the exception of amino acid 341, all of these residues are located in loop regions between regions of antiparallel β-pleated sheets (the *x* and *y* faces) of the protein and are thus exposed to the exterior milieu (38). The amino acid differences in the C_H3 domain are too numerous to define unequivocally a serological determinant. To achieve a final definition of allotypic determinants in molecular terms, a more detailed three-dimensional structure of this molecule may provide clues indicating the proximity of different amino acids to each other. Hybrid molecules, containing parts of two allotypes, might also provide further correlations between serology and primary structure of immunoglobulins.

We thank Drs. M. A. Davies, T. Imanishi-Kari, and K. Rajewsky for stimulating and encouraging discussions, S. Quester for excellent technical assistance, and W. Huelsenbeck for help with some of the experiments. This work was supported by the Deutsche Forschungsgemeinschaft through SFB 74 to B.M.-H. and by grants from the American Cancer Society and the National Cancer Institute to D.B.; A.B. is a Postdoctoral Fellow of the Medical Foundation; D.B. is an American Cancer Society Professor of Microbiology.

1. Lieberman, R. (1978) Springers Seminars in Immunopath. **1**, 7–30.
2. Oudin, J. (1960) *J. Exp. Med.* **112**, 107–124.
3. Shimizu, A., Takahashi, N., Yamawaki-Kataoka, Y., Nishida, Y., Kataoka, T. & Honjo, T. (1981) *Nature (London)* **289**, 149–153.
4. Oi, V. T. & Herzenberg, L. A. (1979) *Mol. Immunol.* **16**, 1005–1017.
5. Seman, F., Dognin, M. J., Stanislawski, M., Seman, M. & Strosberg, A. D. (1979) Immunol. Lett. 1, 141–144.
6. Dognin, M. J., Lauwereys, M. & Strosberg, A. D. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4031–4035.
7. Wilson, G. G. & Murray, N. E. (1979) *J. Mol. Biol.* **132**, 471–491.
8. Murray, N. E., Sandra, A. B. & Murray, K. (1979) *J. Mol. Biol.* **132**, 493–505.
9. Rothstein, R. J., Lau, K. F., Bahl, C. P., Narang, S. A. & Wu, R. (1980) *Methods Enzymol.* **68**, 98–109.
10. Reth, M., Haemmerling, G. J. & Rajewsky, K. (1978) *Eur. J. Immunol.* **8**, 393–400.
11. Bothwell, A. L. M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. & Baltimore, D. (1981) *Cell* **24**, 625–637.
12. Schreier, P. H. & Cortese, R. (1979) *J. Mol. Biol.* **129**, 169–172.
13. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
14. Maat, J. & Smith, A. J. H. (1978) *Nucleic Acids Res.* **5**, 4537–4546.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87**, 107–110.
17. Anderson, S., Gait, M. J., Mayol, L. & Young, I. G. (1980) *Nucleic Acids Res.* **8**, 1731–1743.
18. Staden, R. (1977) *Nucleic Acids Res.* **4**, 4037–4051.
19. Messing, J., Gronenborn, B., Mueller-Hill, B. & Hofschneider, P. H. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
20. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161–178.
21. Fougereau, M., Bourgois, A., de Paval, C., Rocca-Serra, J. & Schiff, C. (1976) Ann. Immuniol. (Inst. Past.) **127C**, 607–631.
22. Sikarov, J. L., Auffray, Ch. & Rougeon, F. (1980) *Nucleic Acids Res.* **8**, 3143–3155.
23. Smith, A. J. H. (1980) *Methods Enzymol.* **65**, 560–580.
24. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639.
25. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.
26. Smith, G. P. (1976) *Science* **191**, 528–535.
27. Slightom, J. L., Blechl, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
28. Egel, R. (1981) *Nature (London)* **290**, 191–192.
29. Miyata, T., Yasunaga, T., Yamawaki-Kataoka, Y., Obata, M. & Honjo, T. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2143–2147.
30. Baltimore, D. (1981) *Cell* **24**, 592–594.
31. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1979) *J. Exp. Med.* **149**, 1299–1313.
32. Kabat, E. A. (1980) *J. Immunol.* **125**, 961–969.
33. Klein, J. (1975) *Biology of the Mouse Histocompatibility-2 Complex* (Springer, New York).
34. Klein, J. (1979) *Science* **202**, 516–521.
35. Nathenson, S. G., Uehara, H., Ewenstein, B. M., Kindt, T. J. & Coligan, J. E. (1981) *Annu. Rev. Biochem.* **50**, 1025–1051.
36. Orr, H. T., Lopez de Castro, J. A., Parham, P., Ploegh, H. L. & Strominger, J. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4395–4399.
37. Burton, D. R., Boyd, J., Brampton, A. D., Easterbrook-Smith, S. B., Emanuel, E. J., Novotny, J., Rademacher, T. W., van Schravendijk, M. R., Sternberg, M. J. E. & Dwek, R. A. (1980) *Nature (London)* **288**, 338–344.
38. Beale, D. & Feinstein, A. (1976) *Q. Rev. Biophys.* **9**, 135–180.