Nucleotide Sequence of a Rabbit IgG Heavy Chain from the Recombinant *F-I* Haplotype

Kenneth E. Bernstein, Cornelius B. Alexander, and Rose G. Mage

Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205

Abstract. We report the sequence of a cDNA encoding a rabbit immunoglobulin γ heavy chain of d12 and e14 allotypes with high homology to partial cDNA sequences from rabbits of d11 and e15 allotypes. The encoded rabbit protein shows homologies with human (68–70%) and mouse (60–63%) γ chains. The nucleotide sequence homologies of the CH domains range from 76–84% with human and 64–76% with mouse sequences. Comparison of the portion of VH encoding amino acid positions 34–112 with a previously determined VH sequence of the same allotype shows high conservation of sequences in the second and third framework segments but more marked differences both in length and encoded amino acids of the second and third complementarity-determining regions (CDRs). We also found a high degree of homology with a human genomic *V*-region, VH26 (77%) and a remarkable similarity between rabbit and human second CDR sequences and human genomic *D* minigenes. These results provide additional evidence that *D* minigene sequences share information with the CDR2 portion of VH regions.

Introduction

Advances in molecular biology have contributed to an in-depth understanding of the organization of immunoglobulin (*Ig*) genes. This understanding continues to supply new insights into mechanism underlying many phenomena previously observed and studied. We have begun to apply the approaches of gene cloning and sequencing to our investigations of genetically controlled alternative forms of rabbit Igs (allotypes; Pavirani et al. 1982, Bernstein et al. 1982, 1983). The rabbit has four allotypic groups for which structural protein data are available (Mage et al. 1973, Kindt 1975): the VHa group that correlates with multiple amino acid differences in the first and third framework (FR) segments of the heavy chain variable region (VH), the d and e allotypes, each correlated with single amino acid differences in the

constant region of γ heavy chains (C_γ) and the b group of kappa (κ) light chain allotypes. The VHa and $C\gamma$ de allotypes are controlled by linked genes, and there are certain prevalent haplotypes so that such combinations as VHa3de12,14 or VHa2de11,15 are rare in or absent from the rabbit populations that have so far been examined (Dubiski and Good 1972, Mage et al. 1982).

This laboratory reported an apparent recombinant (Mage et al. 1971) between the *I* (VHa1,de12,14) and *F* (VHa2,de12,15) haplotypes that resulted in the *F-I* haplotype (VHa2,de12,14; Mage et al. 1971, 1982). Structural studies of the Igs from this animal and its progeny showed that the $C\gamma$ allotype e14 in the recombinant was correlated with the same sequence difference, threonine replacing alanine (Carta-Sorcini et al. 1973) at position 309, that had been found in other e14 rabbits (Appella et al. 1971). In addition, the Fd portion of the heavy chain had an amino acid composition characteristic of the VHa2 allotype (Carta-Sorcini et al. 1973). We recently published the VHa2 sequence of a cDNA encoding a μ heavy chain derived from this haplotype and confirmed that indeed the encoded variable region was characteristic of VHa2 proteins (Bernstein et al. 1982).

We report here the sequence of a cDNA clone encoding part of another VHa2 region and the entire constant region of a γ heavy chain from the recombinant (*F-I*) haplotype. We compare this nucleotide sequence with our first VHa2 sequence, with partial γ chain sequences available from rabbits of d11 and e15 allotypes, and with sequences of human and mouse γ chains.

Materials and Methods

The construction of cDNA clones from splenic mRNA isolated from homozygous rabbits of allotypes VHa2, de 12,14 and b5 infected with *Trypanosoma equiperdum* has been described (Bernstein et al. 1983). We identified clones by colony hybridization by first probing with a ^{32}P-labeled restriction fragment from clone pμ3 containing VHa2 and Cμ sequences (Bernstein et al. 1982). Forty-two positive colonies of the 528 screened were picked and rescreened with a ^{32}P-labeled restriction fragment containing VHa2 sequences and with labeled, sucrose-gradient-fractionated poly (A)$^+$ RNA enriched for mRNAs encoding γ chains (Bernstein et al. 1983). Of the 42 colonies screened, 12 were positive with both probes, and clone pγB1-12,14 was randomly chosen for further study. Plasmid DNA was isolated as described (Ish-Horowicz and Burke 1981). DNA was sequenced as described (Maxam and Gilbert 1980), labelling with either γ^{32}P-ATP and polynucleotide kinase, α^{32}P cordycepin-5′ triphosphate and terminal deoxynucleotidyl transferase, or α^{32}P dCTP and the large fragment of DNA polymerase I to "fill in" 5′ extensions.

Results

Figure 1 shows the strategy used to sequence clone pγB1-12,14. Figure 2 lists the nucleic acid sequence and deduced amino acid sequence and compares the variable region with our previously published pμ3 sequence and the constant region with previously published protein sequences (Hill et al. 1966, 1967, Delaney and Hill 1968, Lebovitz et al. 1968, Prahl et al. 1969, Appella et al. 1971, Pratt and Mole 1975) and partial nucleic acid sequences of rabbit γ chains of d11 and e15 allotypes (Heidmann and Rougeon 1982, Martens et al. 1982). It is immediately clear that clone pγB1-12,14 encodes the entire constant region of rabbit γ chain, a portion of
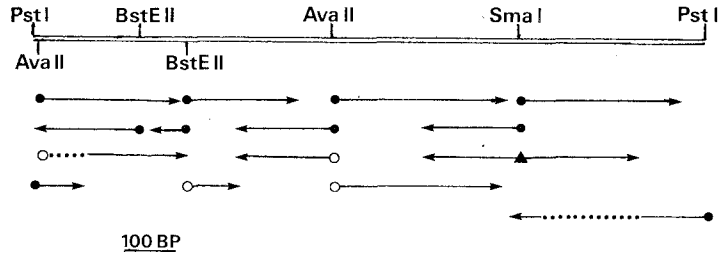
Fig. 1. Strategy of sequencing cloned cDNA pγB1-12,14. *Open circles* indicate fragments labeled by filling in using the large fragment of DNA polymerase I and α^{32}PdCTP; *solid circles* indicate 5' labeled and filled triangle 3' labeled fragments. The *dotted lines* indicate portions of fragments not re-sequenced.

variable region beginning with the codon for the amino acid at position 34, as well as probable *D*- and *J*-encoded regions. The clone also contains 29 bases of the 3' UT region sequence.

Clone pγB1-12,14 was prepared from mRNA of an animal expressing the VHa2 and the de12,14 allotypes. In several ways, clone pγB1-12,14 encodes a more typical *V*-region protein sequence than does pμ3, although the two a2 variable regions are 89.5% homologous for 189 positions encoding amino acids 34–95. Whereas clone pμ3 had codons for Gly and Ser at positions 93 and 94, pγB1–12,14 encodes the more usual Ala and Arg. In the third framework segment, codons for amino acids characteristic of the VHa2 allotype (shown in boxes) are found at each of the seven positions, including (AGA) for Arg 71 that was not found in clone pμ3. The sequence that may be *D*-encoded is homologous to a portion of the *D*-encoded sequence of pμ3 and the *J*-encoded sequences differ by two bases. The VH sequence of pγB1-12,14 is 77% homologous for 195 bases compared with a human genomic *V* region sequence, VH26 (Matthyssens and Rabbits 1980). Like clone pμ3, the second complementarity-determining region (CDR2) has a remarkable degree of homology to the CDR2 of the human *V* gene and to human *D* minigenes (Siebenlist et al. 1981; Fig. 3). There is a stretch of 19 bases in which only two differences between VH26 and pγB1-12,14 occur. Another stretch of six identical bases is separated by six nonhomologous bases, five of which are identical with those in the pμ3 sequence.

In the constant region, as indicated by the boxed amino acids in Figure 2 at positions 225 and 309 (EU Index; Kabat et al. 1983), clone pγB1-12,14 encodes the expected threonine (ACG) at the positions of both the d12 and the e14 allotypes. In the sequences of rabbit γ chains of d11 and e15 allotypes, the (ATG) codon for methionine at position 225 (Prahl et al. 1969, Martens et al. 1982) and the (GCG) codon for alanine at position 309 were found (Appella et al. 1971, Heidmann and Rougeon 1982, Martens et al. 1982). For the published segments of rabbit γ DNA sequences compared with pγB1-12,14 in Figure 2–4 over the 735 bases available for comparison there is 97.6% homology. There is one silent change in the codon for Arg 340 and one replacement change in the codon for Glu or Gln 311 where our sequence agrees with that of Heidmann and Rougeon (1982), but differs from that of Martens and co-workers (1982). Depending upon which sequence is compared, there are 16 or 18 base changes, and 4 or 5 amino acid replacements result. In addition to the expected allotypic differences, we found substitutions of Gln (CAG)

for Glu (GAG) at positions 268 and 311 and Asn (AAC) for Ser (AGC) at position 408. Figure 2 also indicates where the previously determined amino acid sequences differ from that encoded by clone pγB1-12,14. There are seven clarifications or discrepancies of amide assignments for Asn-Asp and five for Gln-Glu. The protein sequence reported by Pratt and Mole (1975) also had two additional amino acids (Pro-Val) between positions 188 and 189 that are not found in clone pγB1-12,14. Two other discrepancies between reported protein sequences and the DNA sequence found by ourselves and others (Heidmann and Rougeon 1982, Martens et al. 1982) appear to be due to typographical errors in the 1967 paper of Hill and co-workers (1967). Earlier and more recent papers from that laboratory (Hill et al. 1966, Lebovitz et al. 1968, Delaney and Hill 1968), as well as the confirmatory work of Appella and co-workers (1971), found Glu rather than Gly at position 380 and Tyr rather than Trp at position 404.

Table 1 shows homology comparisons of the nucleic acid and encoded amino acid sequence of the rabbit γ chain CH1, CH2 and CH3 domains with those of the human γ1, γ2, γ4 and mouse γ1, γ2a and γ2b sequences (Ellison et al. 1982, Ellison and Hood 1982, Obata et al. 1980, Sikorav et al. 1980, Tucker et al. 1979). A three codon gap exists in the CH1 domain of this rabbit γ chain sequence compared with the human and mouse γ chains at positions corresponding to amino acids 193–195 (Eu index; Kabat et al. 1983). The mouse sequences also have deleted a codon corresponding to Ser 177 of our sequence. The CH1 domain of rabbit γ has a similar degree of homology to mouse and human CH1 sequences (Table 1). The rabbit sequence is slightly more homologous to the human than the mouse in CH2 and markedly more homologous to the human in CH3 (80% vs 64–70% nucleotide homology). Similarly, although the rabbit sequence is much less homologous to human or mouse sequences in the hinge region than in the domains, there is somewhat greater homology to the human than to the mouse (Fig. 4). The rabbit γ hinge region is relatively short. Including the hinge regions, the overall protein sequence homologies of the rabbit γ chain to human sequences are 70, 68, and 69% for γ1, γ2 and γ4; to mouse they are 60, 63, and 63% to γ1, γ2a, and γ2b.

Discussion

Clone pγB1-12,14 was constructed from mRNA obtained from rabbits that are direct descendants of a mating in which a recombination event between VHa and Cγ

Fig. 2. The nucleotide sequence and deduced amino acid sequence of pγB1-12,14 compared with the variable region sequence of clone pμ3 (Bernstein et al. 1982), the published constant region protein sequences (Appella et al. 1971, Hill et al. 1966, 1967, Lebovitz et al. 1968, Delaney and Hill 1968, Prahl et al. 1969, Pratt and Mole 1975) and partial nucleotide sequences (Heidmann and Rougeon 1982, Martens et al. 1982). The amino acids that differ in the protein sequences are shown above the pγB1-12,14 sequence. *Lines* signify identities to the nucleotide sequence of pγB1-12,14 and *dashes* indicate gaps introduced to maximize homologies. The nucleotide sequence encoding amino acid positions 208–390 (p2a2) was from a donor rabbit of γ allotypes d11 and e15 (Martens et al. 1982). The nucleotide sequence encoding amino acid positions 268–447 and the 3′ untranslated region was from a donor of e15 and unspecific d allotype (Heidmann and Rougeon 1982). The amino acid numbers are according to the standard system of Kabat et al. (1983) (EU index).

Amino Acid Position 34   40   50
                     MetSerTrpValArgGlnAlaProGlyLysGluLeuGluTrpIleGlyTyrIleSer--Ty-rGlyGlySerAlaTyrTyr
pȳB1-12,14   1 ATGAGCTGGGTCCGACAGGCTCCAGGGAAGGAGCTGGAGTGGATCGGATACATTAGT--TA-TGGTGGTAGTGCATACTAC
                 G——A———————C——————————————G——————————————T————GG——C——————————A———TCTT
pμ3   ValAsn   Gly   GlyThr   Thr   Leu

V_H   60   70   80 82 A B C
               AlaSerTrpAlaLysSerArgSerThrIleThrArgAsnThrAsnGluAsnThrValThrLeuLysMetThrSerLeuThr
         79 GCGAGCTGGGCGAAAAGCCGATCCACCATCACCAGAAACACCAACGAGAACACGGTGACTCTGAAAATGACCAGTCTGACA
               ——————A——————————————————————————————————C——————————————————————
               Asn   Ser

                     90   92   D ?   101   J
               AlaAlaAspThrAlaThrTyrPheCysAlaArg-------HisTrp---------------GlyIleTrpGlyProGly
         160 GCCGCGGACACGGCCACCTATTTCTGTGCGAGA-------CATTGG---------------GGCATCTGGGGCCCAGGC
               ——G——————————————————G———TGGCGCCAAT——————AAAATGAGTTTTTCAAT-C——————————
               Gly   GlySerGlyAlaAsnIleGluAsnGluPhePheAsnAla

D_H, J_H

J_H, CH1   110   CH1 120   130
               ThrLeuValThrValSerSerGlyGlnProLysAlaProSerValPheProLeuAlaProCysCysGlyAspThrPro
         217 ACCCTGGTCACCGTCTCCTCAGGGCAACCTAAGGCTCCATCAGTCTTCCCACTGGCCCCCTGCTGCGGGGACACACCC
               ————————————————————————————————G

            140   150   160
               SerSerValThrLeuGlyCysLeuValLysGlyTyrLeuProGluProValThrValThrTrpAsnSerGlyThrLeu
         295 AGCTCCACGGTGACCCTGGGCTGCCTGGTCAAAGGCTACCTCCCGGAGCCAGTGACCGTGACCTGGAACTCGGGCACCCTC

            Asp   ·   170   180   ProVal
               ThrAsnGlyValArgThrPheProSerValArgGlnSerSerGlyLeuTyrSerLeuSerSerValValSerVal------
         376 ACCAATGGGGTACGCACCTTCCCGTCCGTCCGGCAGTCCTCAGGCCTCTACTCGCTGAGCAGCGTGGTGAGCGTG------

            190 196   200   205
               ThrSerSerGlnProValThrCysAsnValAlaHisProAlaThrAsnThrLysValAspLysThrVal
cDNA(20)   451 ACCTCAAGCAGCCAGCCCGTCACCTGCAACGTGGCCCACCCAGCCACCAACACCAAAGTGGACAAGACCGTT

            220   225 231←————CH2
                               Met
Hinge-CH2   AlaProSerThrCysSerLysProThrCysProProProGluLeuLeuGlyGlyProSerValPheIlePheProProLys
cDNA(20)   523 GCACCCTCGACATGCAGCAAGCCCACGTGCCCACCCCCTGAACTCCTGGGGGGACCGTCTGTCTTCATCTTCCCCCCAAAA
               ——G————G——————————————T——————————C————A———G————————T—T————T—T
                              Met

            250   260   Glu 270
               ProLysAspThrLeuMetIleSerArgThrProGluValThrCysValValValAspValSerGlnAspAspProGluVal
cDNA(20)   604 CCCAAGGACACCCTCATGATCTCACGCACCCCCGAGGTCACATGCGTGGTGGTGGACGTGAGCCAGGATGACCCCGAGGTG
cDNA(19)   ——G
                                                          Glu

                    280   290   300
            Glx   AsxAsx   Asp
            GlnPheThrTrpTyrIleAsnAsnAsnGluGlnValArgThrAlaArgProProLeuArgGluGlnGlnPheAsnSerThrIle
cDNA(20)   685 CAGTTCACATGGTACATAAACAACGAGCAGGTGCGCACCGCCCGGCCGCCGCTACGGGAGCAGCAGTTCAACAGCACGATC
cDNA(19)

                          310   320
                         Ala   Glu   Asp
            ArgValValSerThrLeuProIleThrHisGlnAspTrpLeuArgGlyLysGluPheLysCysLysValHisAsnLysAla
cDNA(20)   766 CGCGTGGTCAGCACCCTCCCCATCACGCACCAGGACTGGCTGAGGGGCAAGGAGTTCAAGTGCAAAGTCCACAACAAGGCA
cDNA(19)   ——G————G
                                Glu
                           Ala

            330   340 CH3Glu   350
               LeuProAlaProIleGluLysThrIleSerLysAlaArgGlyGlnProLeuGluProLysValTyrThrMetGly
CH2-CH3   847 CTCCCGGCCCCCATCGAGAAAACCATCTCCAAAGCCAGAGGGCAGCCCCTGGAGCCGAAGGTCTACACCATGGGC
cDNA(20)
cDNA(19)   ——————————————————————G

                       360   370
            Gln   Asp
             ProProArgGluGluLeuSerSerArgSerValSerLeuThrCysMetIleAsnGlyPheTyrProSerAspIleSerVal
         922 CCTCCCCGGGAGGAGCTGAGCAGCAGGTCGGTCAGCCTGACCTGCATGATCAACGGCTTCTACCCTTCCGACATCTCGGTG

          380   390   400
          (Gly)   Asp   (Trp)
          GluTrpGluLysAsnGlyLysAlaGluAspAsnTyrLysThrThrProAlaValLeuAspSerAspGlySerTyrPheLeuTyr
cDNA(20) 1003 GAGTGGGAGAAGAACGGGAAGGCAGAGGACAACTACAAGACCACGCCGGCCGTGCTGGACAGCGACGGCTCCTACTTCCTCTAC
cDNA(19)

          Ser   410   420   430
          AsnLysLeuSerValProThrSerGluTrpGlnArgGlyAspValPheThrCysSerValMetHisGluAlaLeuHisAsn
cDNA(19) 1087 AACAAGCTCTCAGTGCCCACGAGTGAGTGGCAGCGGGGCGACGTCTTCACCTGCTCCGTGATGCACGAGGCCTTGCACAAC
           —G——
          Ser

            440   446
          HisTyrThrGlnLysSerIleSerArgSerProGlyLysTrm
CH3-3'UT 1168 CACTACACGCAGAAGTCCATCTCCCGCTCTCCGGGTAAATGAGCGCTGTGCCGGCGAGCTGCCCCTCT
cDNA(19)

```
Amino Acid Position    49                          59 49                                      59
         Human D2    AGGA--TATT----GTAGTGGTGGTAGCTGCTACTCC
                     III     III      II IIIIIIIIII     IIII I
   Rabbit pγB1-12,14 GGATACATTAGT--TA-TGGTGGTAGTGCATACTAC Gly Tyr Ile Ser --T y-r Gly Gly Ser Ala Tyr Tyr
                     I    IIIIII  II IIIIIIIII  IIIIIIII
         Human V_H26 TCAGCTATTAGTGGTAGTGGTGGTAGCACATACTAC Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr
                     I   IIII   IIII IIII    IIIII
   Rabbit pμ3        GGATATATT---GGTACTGGT------ACATATCTT Gly Tyr Ile --- Gly Thr Gly --- --- Thr Tyr Leu
                     III  IIII    IIIIIIIII       III
         Human D1    AGGA--TATT----GTACTGGTGGTGTATGCTATACC
```

Fig. 3. The nucleotide sequences of part of the second complementarity-determining region, CDR2, (codons for Gly 49 through Tyr 59 of pγB1-12,14) of two rabbit and one human V_H region (Matthyssens and Rabbits 1980) compared with the sequences of two human D minigenes (Siebenlist et al. 1981). *Vertical lines* indicate identities and *dashes* indicate gaps introduced to maximize homologies.

Table 1. Comparison of the homologies* of rabbit γ heavy chain sequences to mouse and human γ chains

Rabbit pγB1-12,14	CH1 %[†]	CH2 —[†]	CH3 —[†]
Human			
γ1	76	84	81
	(66)	(74)	(72)
γ2	76	83	80
	(68)	(70)	(71)
γ4	76	82	79
	(69)	(74)	(69)
Mouse			
γ1	75	75	70
	(65)	(64)	(55)
γ2a[a]	76	75	64
	(70)	(68)	(55)
γ2b[a]	77	75	64
	(65)	(68)	(59)

* Nucleic acid sequence homologies are given above protein sequence homologies shown in parentheses.
[†] Percentages are expressed as 100 minus (the number of positions where a difference occurs/total number or positions compared) times 100. Gaps are scored as one difference at a single position.

genes was observed (Mage et al. 1971). The result of the crossover between the F (VHa2,de12,15) and I (VHa1,de12,14) haplotypes was the F-I haplotype with VHa2 and de12,14. The point of crossover is postulated to have been between the VHa and Cγe-encoding DNA but has not been localized further (Mage et al. 1982). The high degree of homology of the γ chain sequences from animals of the d11 and e15 allotypes and our sequence makes it difficult to determine whether this crossover affected the structure of the γ chain. Additionally, since only partial nucleotide sequences of γ chains of d11 and e15 allotypes have been published, we were only able to compare 735 bases of which only 24 were from CH1. The comparison showed 16 or 18 base changes four or five of which resulted in amino acid replacements including the two expected allotypic positions. With the exception of the Ser to Asn replacement at position 408, all the differences were in the hinge region and CH2 domain with 10 of the 18 differences within 100 bases centered on the d position and 5 of the 18 clustered within a 71-base segment encompassing the e position. This may indicate that the allotypic variation arose through nucleic acid

```
Nucleotide number  523                                  555
Rabbit pyB1-12,14  GCACCCTCGACATGCAGCAAGCCC---ACGTGCCCA
G1HU               --AG-----AAATCT---TGA-----AA-TCAC--A---------CCGTGC
G2HU               --AG--G-AAATGT---TGT--------------------CCGTGCCCA
G4HU               --AGT---AAATATGGT-----------------C-A---------TCATGCCCA
G1MOUSE            --TG-----AG-GAT---TG-TTGT-------A-CCTTGCATATGTACA
G2aMOUSE           --AG-----AGAGGGCC---CA-TC-------A-CC-TGTCCTCCATGC-----------AAATGCCCA
G2bMOUSE           --AG-----AGCGGGCC---TTTCAA-AATC-ACCC-TGTCCTCCATGCAAGGAGTGTCACAAATGCCCA
```

Fig. 4. The nucleotide sequence of the portion of pyB1-12,14 cDNA encoding the hinge region compared with those of human γ1 (Ellison et al. 1982), γ2 (Ellison and Hood 1982), γ4 (Ellison and Hood 1982) and mouse γ1 (Obata et al. 1980), γ2a (Sikorav et al. 1980), and γ2b (Tucker et al. 1979) sequences. *Lines* signify identity to the rabbit sequence and *dashes* indicate gaps introduced to maximize homologies.

sequence changes in permitted portions of the protein or nucleotide sequence. Marked differences in the degree of sequence divergence have also been observed when comparisons between allotypic or isotypic forms of mouse γ chains were made domain by domain (Miyata et al. 1980, Schreier et al. 1981, Ollo and Rougeon 1982).

The hinge region has been observed to vary more between different mouse and human γ chains than the domains. Thus it is not surprising that the rabbit sequence is least homologous to the mouse and human γ chains in this region (Fig. 4). On the other hand, the hinge region sequence of clone pyB1–12,14 and the previously published sequence from d11 allotype are very homologous with two silent changes in addition to the expected allotype difference at position 225.

In hares and cottontails, closely related genera (*Lepus* and *Sylvilagus*) of the family *Leporidae*, the d11 and d12 allotypes are not detected (Aggarwal and Mandy 1976). This is consistent with the sequence differences found at the protein level, where instead of the Met or Thr at position 225 that rabbits possess, hares have Pro and cottontails have Leu. Similarly, the related genera have amino acid replacements in the vicinity of the e14-e15-correlated position 309. *Lepus timidus* have Lys replacing Ala or Thr at 309 and thus no e15 antigenic reactivity (Teherani et al. 1979). In some hares Thr 305 is replaced with Ile and Arg 315 is replaced with Ser; in Pikas the Arg 315 is replaced with Lys. These changes alter the e15 antigenic determinant (Teherani et al. 1979). A similar fine structural determinant may be influenced by position 311. Our e14 sequence and the e15 sequence of Heidmann and Rougeon (1982) both encode Gln at this position; Martens and co-workers (1982) found the Glu previously identified by amino acid sequencing. This Gln-Glu difference at 311 may lead to different fine specificity of the e-group determinants in different rabbits.

In comparing our deduced amino acid sequence of the CH1 domain with the available amino acid sequence, we find an insertion of Pro and Val between positions 188 and 189 (Fig. 2). This was a difficult area for protein sequencing and the paper of Pratt and Mole (1975) corrected an earlier sequence of Fruchter and co-workers (1970) from the same laboratory. It remains possible that the corrected sequence was still in error. Alternatively, the difference could be due to the different genetic source of the γ chains, since they were probably γ chains of the prevalent de12,15 allotype. If the CH1 sequence difference is a real one, it would not be involved in the serologically detectable e allotypic determinant found on isolated Fc fragments of rabbit IgG (CH2 and CH3 domains). It is unlikely that the sequence difference at position 408 in the CH3 domain contributes to an antigenic difference between e14 and e15. The homologous position in the crystallographic structure of

human Fc is in the interior CH3 domain interface and the Ser to Asn replacement would result in only a small volume increase (Deisenhofer 1981). The interchange of a Ser in e15 for an Asn in e14 may have been overlooked when Appella and colleagues (1971) compared the compositions of tryptic peptides from Fc fragments of e15 and e14 IgGs. Tryptic peptide T5 is 17 amino acids long and was obtained in low yields from both e15 and e14 IgGs (12% and 17%) because it is difficult to purify. The compositions of both peptides corresponded to that expected from the previously reported sequence of e15 IgG, which has 2 Asp and 3 Ser. The e15 peptide had 1.6 Asp and 3.4 Ser vs 2.3 Asp and 2.9 Ser in the e14. The fractional recoveries of Ser and Asp residues could mean that both sequences were represented in IgGs from both allotypes or could have been due to the experimental imprecision in measuring these amino acids. Additional DNA sequences from animals of the two genetic types may eventually allow distinction between these possibilities.

When compared with the first rabbit V-region DNA sequence published, (pμ3), the VH and J regions encoded by clone pγB1-12,14 are highly homologous but the two sequences are very different in the D-encoded segment. Whereas the D region of clone pμ3 is quite long, that of clone pγB1-12,14 is quite short. Following Cys, Ala, Arg at positions 92–94 and preceding Gly, Ile at positions 101 and 102 there are only two codons for the amino acids His and Trp. Whether these are encoded by a small D section or by a J region cannot be determined. However, examples of Igs with short D-encoded regions have been identified among the mouse anti-inulin myeloma proteins (Vrana et al. 1978). The number of JH genes present in the rabbit genome is not known but protein sequencing has delineated a prototypic sequence between positions 101 and 113 (Kabat et al. 1983). Our previously published pμ3 sequence encoded Ala, Ile at positions 101 and 102 rather than the more usual Asp, Val; clone pγB1-12,14 had Gly, Ile at comparable positions and like clone pμ3 is identical with the prototype for the remainder of J. Although these V regions are both typical of VHa2 allotype, it is unlikely that Ile at position 102 is allotype-specific since Ser, Ile was found at positions 101 and 102 in an a1 VH region sequence (protein 723569; James and Freedman 1977).

Comparison of the second and third framework regions of pγB1-12,14 and pμ3 shows very highly homology: six changes in 141 bases corresponding to amino acid positions 36–49 and 65–94. Thus it appears that rabbits of a particular allotype can have extraordinary conservation of large areas of sequence within the "variable" region. How it is that such conservation was maintained and yet variability generated can be glimpsed by examining the CDR2 and CDR3 of these clones. The CDRs of the two sequences are strikingly different both in length and in encoded amino acid sequences and thus can be expected to contribute to combining sites with very different structures. As previously noted, despite marked differences in the encoded protein structures, there is remarkable nucleotide sequence homology between the CDR2 of clone pμ3 and that of a human genomic clone VH26. In addition, homology exists between the sequence of the CDR2 of pμ3 and sequences of human genomic D minigenes. The CDR2 region of clone pγB1-12,14 also has homology to these two minigenes as well as remarkable homology to the CDR2 of VH26. As shown in Figure 3, 17 of 19 bases and over a longer distance 25 of 30 positions can be aligned. Moreover, there are nine contiguous bases that match the human $D2$ minigene (Siebenlist et al. 1981) within the same region where two human

VH sequences, VH26 (Matthyssens and Rabbits 1980) and HG3 (Rechavi et al. 1983), were found to have 14 and 13 nucleotide matches (Wu and Kabat 1982). The protein sequence at amino acid positions 54–56 (Gly, Gly, Ser) is found in VH26, HG3, pγB1–12,14 and two fractions (3381 and 3381-2; Haber et al. 1977) of one rabbit's type III-specific pneumococcus antibody. Why there should be this homology between different species as well as between rabbit *V* regions in an area of hypervariability is difficult to understand but this second example in the rabbit system indicates that the phenomenon is not unique to any one molecule. Perhaps this is a vestige from the time when a primitive *V* region was assembled from primordial shorter subunits (Ohno et al. 1982) and CDR1 and CDR2 formed from minigenes similar to the way CDR3 is today. Even today, the operation of gene conversion (Baltimore 1981, Egel 1981) or minigene insertion mechanisms has not been ruled out (Wu and Kabat 1970, 1982, Kabat et al. 1978).

References

Aggarwal, S. and Mandy, W. J.: Lagomorph IgG hinge region: allotype associated amino acid sequence variations. *Immunochemistry 13:* 215–220, 1976

Appella, E., Chersi, A., Mage, R. G., and Dubiski, S.: Structural basis of the A14 and A15 allotypic specificities in rabbit immunoglobulin G. *Proc. Natl. Acad. Sci. U.S.A. 68:* 1341–1345, 1971

Baltimore, D.: Gene conversion: some implications for immunoglobulin genes. *Cell 24:* 592–594, 1981

Bernstein, K. E., Reddy, E. P., Alexander, C. B., and Mage, R. G.: A cDNA sequence encoding a rabbit heavy chain variable region of the VHa2 allotype showing homologies with human heavy chain sequences. *Nature 300:* 74–76, 1982

Bernstein, K. E., Pavirani, A., Alexander, G., Jacobsen, F., Fitzmaurice, L., and Mage, R.: Use of *Trypanosoma equiperdum* infected rabbits as a source of splenic mRNA: construction of cDNA clones and identification of a rabbit μ heavy chain clone. *Mol. Immunol. 20:* 89–99, 1983

Carta-Sorcini, M., Appella, E., Inman, J. K., and Mage, R.: The immunoglobulins derived from a VH-CH recombinant rabbit and its normal relatives. II. Chemical studies of allotype related differences in Fd and Fc fragments. *Immunochemistry 10:* 449–453, 1973

Deisenhofer, J.: Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9. and $2.8 - A°$ resolution. *Biochemistry 20:* 2361–2370, 1981

Delaney, R. and Hill, R. L.: The amino acid sequence of the Fc fragment of rabbit immunoglobulin G. II. Cleavage of the Fc fragment with cyanogen bromide. *J. Biol. Chem. 243:* 4206–4215, 1968

Dubiski, S. and Good, P. W., Jr.: Population genetics of the heavy chain immunoglobulin allotypes in the rabbit. *Proc. Soc. Exp. Biol. Med. 141:* 486–489, 1972

Egel, R.: Intergenic conversion and reiterated genes. *Nature 290:* 191–192, 1981

Ellison, J. and Hood, L.: Linkage and sequence homology of two human γ heavy chain constant region genes. *Proc. Natl. Acad. Sci. U.S.A. 79:* 1984–1988, 1982

Ellison, J. W., Berson, B. J., and Hood, L. E.: The nucleotide sequence of a human $C_{γ1}$ gene. *Nucleic Acids Res. 10:* 4071–4079, 1982

Fruchter, R. G., Jackson, S. A., Mole, L. E., and Porter, R. R.: Sequence studies of the Fd. section of the heavy chain of rabbit immunoglobulin G. *Biochem. J. 116:* 249, 1970

Haber, E., Margolies, M. N., and Cannon, L. E.: Origins of antibody diversity: insights gained from

amino acid sequence studies of elicited antibodies. *Cold Spring Harbor Symp. Quant. Biol. 41:* 647–659, 1977

Heidmann, O. and Rougeon, F.: Molecular cloning of rabbit γ heavy chain mRNA. *Nucleic Acids Res. 10:* 1535–1545, 1982

Hill, R. L., Delaney, R., Lebovitz, H. E., and Fellows, R. E.: Studies on the amino acid sequence of heavy chains from rabbit immunoglobulin G. *Proc. Royal Soc. Lond. 166:* 159–175, 1966

Hill, R. L., Lebovitz, H. E., Fellows, R. E., Jr., and Delaney, R.: The evolution of immunoglobulins as reflected by amino acid studies of rabbit Fc fragment *In* J. Killander (ed.): *Gamma Globulins, Structure and Control of Biosynthesis*, pp. 109–127, Interscience Publishers, New York, 1967

Ish-Horowicz, D. and Burke, J.: Rapid and efficient cosmid cloning. *Nucleic Acids Res. 9:* 2989–2998, 1981

James, O. and Freedman, M.: Studies of hyperimmune restricted and partially restricted anti-pneumococcal polysaccharide antibodies from allotype-defined pedigreed rabbits. V. Variable region heavy chain sequence analysis of the cyanogen bromide C1 fragment obtained from an unusual restricted anti-SVIII antibody from a homozygous a^1 partially inbred rabbit. *Immunochemistry 14:* 15–24, 1977

Kabat, E. A., Wu, T. T., and Bilofsky, H.: Variable region genes for the immunoglobulin framework are assembled from small segments of DNA-a hypothesis. *Proc. Natl. Acad. Sci. U.S.A. 75:* 2429–2433, 1978

Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M., and Perry, H.: Sequences of proteins of immunological interest: tabulation and analysis of amino acid and nucleic acid sequences of precursors, V-regions. C-regions, J-chain, β_2-microglobulins, major histocompatibility antigens, Thy-1, complement, c-reactive protein, thymopoietin, post-gamma globulin, and α2-macroglobulin. US DHHS, PHS, NIH. 173–193, 1983

Kindt, T. J.: Rabbit immunoglobulin allotypes: structure, immunology, and genetics. *Adv. Immunol. 21:* 35–86, 1975

Lebovitz, H. E., Delaney, R., Fellows, R. E., Jr., and Hill, R. L.: The amino acid sequence of the Fc fragment of rabbit immunoglobulin G. I. The isolation and amino acid composition of the tryptic peptides. *J. Biol. Chem. 243:* 4197–4205, 1968

Mage, R. G., Young-Cooper, G. O., and Alexander, C.: Genetic control of variable and constant regions of immunoglobulin heavy chains. *Nature, New Biology 230:* 63–64, 1971

Mage, R., Lieberman, R., Potter, M., and Terry, W. D.: *In* M. Sela (ed.): Immunoglobulin allotypes. *The Antigens*, pp. 299–376, Academic Press, New York, 1973

Mage, R., Dray, S., Gilman-Sachs, A., Hamers-Casterman, C., Hamers, R., Hanly, W., Kindt, T., Knight, K., Mandy, W., and Naessens, J.: Rabbit heavy chain haplotypes-allotypic determinants expressed by V_H-C_H recombinants. *Immunogenetics 15:* 287–297, 1982

Martens, C. L., Moore, K. W., Steinmetz, M., Hood, L., and Knight, K. L.: Heavy chain genes of rabbit IgG: isolation of a cDNA encoding γ heavy chain and identification of two genomic C_γ genes. *Proc. Natl. Acad. Sci. U.S.A. 79:* 6018–6022, 1982

Matthyssens, G. and Rabbits, T. H.: Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region. *Proc. Natl. Acad. Sci. U.S.A. 77:* 6561–6565, 1980

Maxam, A. M. and Gilbert, W.: Sequencing end-labeled cDNA with base-specific chemical cleavages. *Methods Enzymol. 65:* 498–560, 1980

Miyata, T., Yasunaga, T., Yamawaki-Kataoka, Y., Obata, M., and Honjo, T.: Nucleotide sequence divergence of mouse immunoglobulin γ1 and γ2a genes and the hypothesis of intervening-sequence mediated domain transfer. *Proc. Natl. Acad. Sci. U.S.A. 77:* 2143–2147, 1980

Obata, M., Yamawaki-Kataoka, Y., Takahashi, N., Kataoka, T., Shimizu, A., Mano, Y., Seidman, J. G., Peterlin, B. M., Leder, P., and Honjo, T.: Immunoglobulin γ1 heavy chain gene: structural gene sequences cloned in a bacterial plasmid. *Gene 9:* 87–97, 1980

Ohno, S., Kato, K., Hozumi, T., and Matsunaga, T.: Mouse immunoglobulin coding sequences for the heavy-chain variable region arose as repeats of the two short building blocks. *Proc. Natl. Acad. Sci. U.S.A. 79:* 132–136, 1982

Ollo, R. and Rougeon, F.: Mouse immunoglobulin allotypes: post-duplication divergence of γ2a and γ2b genes. *Nature 296:* 761–763, 1982

Pavirani, A., Mage, R., Jacobsen, F., Reddy, E. P., Bernstein, K., and Fitzmaurice, L.: Construction of a partial rabbit spleen cDNA library and identification of immunoglobulin clones. *Eur. J. Immunol. 12:* 854–860, 1982

Prahl, J. W., Mandy, W. J., and Todd, C. W.: The molecular determinants of the A11 and A12 allotypic specificities in rabbit immunoglobulin. *Biochemistry 8:* 4935–4940, 1969

Pratt, D. M. and Mole, L. E.: Sequence studies of the constant region of the Fd sections of rabbit immunoglobulin G of different allotype. *Biochem. J. 151:* 337–349, 1975

Rechavi, G., Ram, D., Glazer, L., Zakut, R., and Givol, D.: Evolutionary aspects of immunoglobulin V_H gene subgroups. *Proc. Natl. Acad. Sci. U.S.A. 80:* 855–859, 1983

Schreier, P. H., Bothwell, A. L. M., Mueller-Hill, B., and Baltimore, D.: Multiple differences between nucleic acid sequences of the IgG2^a and IgG2^b alleles of the mouse. *Proc. Natl. Acad. Sci. U.S.A. 78:* 4495–4499, 1981

Siebenlist, U., Ravetch, J. V., Korsmeyer, S., Waldmann, T., and Leder, P.: Human immunoglobulin D segments encoded in tandem multigenic families. *Nature 294:* 631–635, 1981

Sikorav, J. L., Auffray, C., and Rougeon, F.: Structure of the constant and 3' untranslated regions of the murine Balb/c γ2a heavy chain messenger RNA. *Nucleic Acids Res. 8:* 3143–3155, 1980

Teherani, J., Capra, J. D., Aggarwal, S., and Mandy, W. J.: Amino acid sequence analysis of group e allotype-related peptides derived from lagomorph IgG. *Eur. J. Immunol. 9:* 690–695, 1979

Tucker, P. W., Marcu, K. B., Slighton, J. L., and Blattner, F. R.: Structure of the constant and 3' untranslated regions of the murine γ2b heavy chain messenger RNA. *Science 206:* 1299–1303, 1979

Vrana, M., Rudikoff, S., and Potter, M.: Sequence variation among heavy chains from inulin-binding myeloma proteins. *Proc. Natl. Acad. Sci. U.S.A. 75:* 1957–1961, 1978

Wu, T. T. and Kabat, E. A.: An analysis of the sequences of the variable regions of Bence Jones proteins myeloma light chains and their implications for antibody complementarity. *J. Exp. Med. 132:* 211–250, 1970

Wu, T. T. and Kabat, E. A.: Fourteen nucleotides in the second complementarity-determining region of a human heavy-chain variable region gene are identical with a sequence in a human D minigene. *Proc. Natl. Acad. Sci. U.S.A. 79:* 5031–5032, 1982