J Mol Evol (1999) 49:23–26

JOURNAL OF MOLECULAR EVOLUTION

Enhanced Evolvability in Immunoglobulin V Genes Under Somatic Hypermutation

Lindsay G. Cowell,[1] Hye-Jung Kim,[2] Tiina Humaljoki,[2] Claudia Berek,[2] Thomas B. Kepler[1]

[1] Biomathematics Graduate Program, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA
[2] Deutsches Rheuma ForschungsZentrum Berlin, Monbijoustrasse 2A, 10117 Berlin, Germany 49 30 2802 6256

Abstract. Darwinian theory requires that mutations be produced in a nonanticipatory manner; it is nonetheless consistent to suggest that mutations that have repeatedly led to nonviable phenotypes would be introduced less frequently than others—if under appropriate genetic control. Immunoglobulins produced during infection acquire point mutations that are subsequently selected for improved binding to the eliciting antigen. We and others have speculated that an enhancement of mutability in the complementarity-determining regions (CDR; where mutations have a greater chance of being advantageous) and/or decrement of mutability in the framework regions (FR; where mutations are more likely to be lethal) may be accomplished by differential codon usage in concert with the known sequence specificity of the hypermutation mechanism. We have examined 115 nonproductively rearranged human Ig sequences. The mutation patterns in these unexpressed genes are unselected and therefore directly reflect inherent mutation biases. Using a χ^2 test, we have shown that the number of mutations in the CDRs is significantly higher than the number of mutations found in the FRs, providing direct evidence for the hypothesis that mutations are preferentially targeted into the CDRs.

Key words: Genome evolution — Adaptability — Somatic hypermutation — Affinity maturation

Correspondence to: Thomas B. Kepler; *e-mail:* kepler@unity.ncsu.edu

Introduction

Darwinian theory requires that mutations be produced in a nonanticipatory manner; it is nonetheless consistent to suggest that mutations that have repeatedly led to nonviable phenotypes would be introduced less frequently than others—if under appropriate genetic control. Many organisms that persist in fluctuating environments depend on genetic variability for their continuing survival. It is of some interest to investigate the possibility of a second-order "evolution of evolvability" that would adapt these organisms to the particular patterns of variability in the environment. The affinity maturation by somatic hypermutation and selection of immunoglobulins (Ig) (Berek and Milstein 1987) provides a simplified yet natural system for the study of just such a process. We have examined 471 mutations occurring in 115 unexpressed human Ig V genes and present here the first conclusive, direct evidence that Ig V genes have evolved an enhanced evolvability under somatic hypermutation.

Immunoglobulin molecules are assembled from two heavy and two light polypeptide chains [for general immunological background, see Paul (1993)]. These heavy and light chains are each encoded by distinct gene segments: a constant-region gene and a variable-region gene (Tonegawa 1983). The variable-region gene itself is constructed by DNA rearrangement involving two (V and J, for light chains) or three (V, D, and J, for heavy chains) minigene segments chosen in a quasi-stochastic process from minigene libraries (Tonegawa 1983). During the immune response, Ig molecules are synthesized by B lymphocytes, utilized by them as surface-bound signal

transduction elements, and secreted by them as effector molecules. In the first few days following infection, a subset of B cells migrates to the lymphoid tissues, where they form germinal centers, the site of affinity maturation (Berek et al. 1991), where the B cells proliferate rapidly and experience somatic hypermutation targeted specifically at the rearranged Ig variable region gene (Lebecque and Gearhart 1990) at rates about 10^6 times above the background genomic rate (McKean et al. 1984). Those B cells whose mutated Ig genes have a higher affinity for the eliciting antigen are selected on this basis and are preferentially expanded (MacLennan 1994); these cells go on to form the memory cell pool, which largely mediates the response during repeated encounters with the same antigen (Berek and Milstein 1987).

Ig have immunoglobulin-fold structures consisting of β-pleated sheets alternating with loops. The β sheets, and other conserved residues necessary for structural integrity, are coded for in the framework regions (FR) of the V gene; the complementarily-determining regions (CDR) encode residues directly involved in antigen binding. Mutations in the FR are more likely to be deleterious, while mutations in the CDR are most likely to alter the affinity (Shlomchik et al. 1989). Indeed, Ig sequences recovered late in the immune response show an enhancement of mutations in the CDR over FR (Berek and Milstein 1987). Thus, if possible, a focusing of mutations into the CDR would improve the efficiency with which B cells adapt to different pathogens.

The hypermutation mechanism displays strong sequence specificity, in that particular nucleotide motifs promote mutation ("hot" motifs), while others have below-average mutability ("cold" motifs) (Betz et al. 1993; Smith et al. 1996). Differential codon usage between CDR and FR may therefore be the mechanism by which the mutability of CDR is enhanced relative to that of FR (Reynaud et al. 1995).

Suggestive but inconclusive evidence for this intriguing hypothesis—based on apparent clustering of mutations in CDR—has been offered (Motoyama et al. 1991; Betz et al. 1993; Dörner et al. 1997). Further suggestion comes from the observed segregation of "hot" (AGY) and "cold" (TCN) serine codons between CDR and FR (Wagner et al. 1995), but this methodology, which ignores correlations among related genes, is statistically problematic. When all codons are analyzed and correlation difficulties obviated, one finds a significant differential codon usage (Kepler 1997). Our purpose in this brief communication is to corroborate this indirect evidence with direct observation of the relative mutation frequencies in the CDR and FR of Ig genes that have not been subject to selection.

Such unselected genes are available for study because of the randomness of the rearrangement process: joining in any of the three reading frames is possible, but only one produces a productive gene (Tonegawa 1983). If the initial rearrangement is out of frame, the cell can still be "rescued" by producing a productive rearrangement on its other chromosome (Tonegawa 1983). This cell will then have, in addition to the productive gene that encodes the Ig molecule the cell produces, a nonproductive gene that experiences somatic hypermutation but cannot be the object of selection since it cannot make a protein. By recovering many such nonproductive genes and analyzing the distribution of mutations they contain, we are able to show conclusively that CDR are inherently more mutable than FR.

Materials and Methods

We have collected over 600 Ig sequences from arthritis patients and healthy donors using DNA isolated from peripheral blood lymphocytes and synovial infiltrate B cells. Peripheral blood mononuclear cells (PBMC) were isolated from the total blood of three rheumatoid arthritis patients and two healthy donors using Ficoll. The PBMC were lysed with proteinase K, and the immunoglobulin genes were amplified in a polymerase chain reaction (PCR) using a primer mix designed to bind all V gene families. The following sequences of oligonucleotides, from 5′ to 3′, were used as primers for the amplification of rearranged Ig genes: VH2, CAGATCACCTTGAAGGAGTCTGG; VH3, GAGGTGCAGCTGGTGGAGTCTGG; VH4, CAGGTGCAGCTGCAGGAGTCGGG; VH6, TACAGCTGCAGCAGTCAGGTCCAGG; JH, CTCACCTGAGGAGACGGTGACC; Vk1, GACATCCAGATGACCCAGTCTCC; Vk2, GATGTTGTGATGACTCAGTCTCC; Vk3, GAAATTGTGTTGACGCAGTCTCC; Vk4, GACATCGTGATGACCCAGTCTCC; Vk5, GGAAATTGTGCTGACTCAGTCTCC; Vk6, GAAATTGTGCTGACTCAGTCTCC; Jk1–4, TACTTACGTTTGATCTCCASCTTG; Jk5, GCTTACGTTTAATCTCCAGTCGTG; V11, CAGTCTGTGTTGACGCAGCCGCC; V12a, CAGTCTGCCCTAGACTCAGCCTGC; V12b, CAGTCTGCCTGATTCAGCCTCC; V13a, TCCTATGTGCTGACTCAGCCACC; V13b, TCTTCTGAGCTGACTCAGGACCC; V14, CACGTTATACTGACTCAACCGCC; V15, CAGGCTGTGCTCACTCAGCCGTC; V16, AATTTTATGCTGACTCAGCCCCA; V17–8, CAGACTGTGGTGACGCAGGAGCC; J12, GAGAGCCACTTACCTAGGACGG; and J13, AGAAGAGACTCACCTAGGACGG.

The PCR was carried out in a 50-μl reaction volume containing 10× buffer and 2U Taq DNA polymerase (Promega); 10 pmol of each primer; a 200 μ*M* concentration each of dATP, dCTP, dTTP, and dGTP; and 0.1 m*M* magnesium chloride. The thermal cycler program was as follows: 2 min at 95°C for one denaturing cycle; 1 min at 94°C, 90 sec at 70°C (for heavy chains) or 65°C (for light chains), 2 min at 72°C for 35 amplification cycles; and 15 min at 72°C for one terminal cycle. The PCR product was electrophoretically separated (visualized by UV light in an ethidium bromide-stained 2% low-melting agarose gel), and product of the correct size (approximately 350 bp) was cut out of the gel and reamplified. In the second PCR, all reagents and concentrations were as in the first PCR except that nested 3′ primers were used. The actual sequences are as follows: JH1, TGAGGAGACGGTGACCAGGGTGCC; JH3, TGAAGAGACGGTGACCATTGTCCC; JH4, TGAGGAGACGGTGACCAGGGTTCC; JH6, TGAGGAGACGGTGACCGTGGTCCC; Jk1, ACGTTTGATTTCCACCTTGGTCCC; Jk2, ACGTTTGATCTCCAGCTTGGTCCC; Jk3, ACGTTTGATATCCACTTTGGTCCC; Jk4, ACGTTTGATCTCCACCTTGGTCCC; Jk5, ACGTTTAATCTCCAGTCGTGTCCC; J12, ACCTAGGACGGTGACCTTGGTCCC; J13, ACCTAGGACGGTCAGCTTGGTCCC; and J15, ACCTAAAACGGTGAGCTGGGTCCC.

The first 95°C denaturing cycle was followed by 39 amplification cycles at the same temperatures listed above and one 72°C 15-min termination cycle. Second-round PCR product of the correct size was

Table 1. A comparison of the Mutation Frequency of CDR and FR

	CDR		FR		
Locus	Mutations/bases	Mutation frequency	Mutations/bases	Mutation frequency	*p*
H	25/629	0.0397	37/1,853	0.0200	0.009
κ	125/6,956	0.0180	200/16,016	0.0125	0.002
λ	42/1,595	0.0263	42/3,300	0.0127	0.001
			Serine codons omitted		
H	15/518	0.0290	34/1,601	0.0212	0.396
κ	88/5,297	0.0166	181/13,793	0.0131	0.078
λ	31/1,190	0.0261	36/2,814	0.0128	0.002

isolated as described above, extracted from the agarose gel using the JETSORB kit (Genomed), and cloned using the ONE STEP TA-CLONING kit (Invitrogen). Plasmid DNA was isolated from clones using phenol/chloroform extraction, and the Ig genes were sequenced using the dideoxy sequencing method with Taq polymerase. Digoxigenin-labeled −40 and +40 primers for the cloning vector (pCR 2.1, One Step kit) were used as sequencing primers.

In patients with rheumatoid arthritis (RA) or reactive arthritis (ReA), lymphocyte infiltrates that resemble germinal centers are found in the inflamed synovial tissue (Schröder et al. 1996; Berek and Schröder 1997). Analysis of the V regions from B cells isolated from these lymphocyte infiltrates indicate that clonal expansion and diversification of the B-cell repertoire take place in the synovial infiltrate (Schröder et al. 1996; Berek and Schröder 1997; Berek and Kim 1997). Cryostat sections were prepared from the synovial tissue of 1 ReA and 3 RA patients as previously described (Schröder et al. 1996). Lymphocyte infiltrates were located by staining with anti-CD20 and plasma cell and follicular dendritic cell specific antibodies (Schröder et al. 1996). B cells were directly isolated from the tissue sections by microdissection, and their rearranged V genes cloned and sequenced as described above.

The germline sequences for all Ig genes sequenced were identified using the international ImMunoGeneTics database (IMGT; Lefranc 1997). Base differences between the Ig genes sequenced and the corresponding germlines were scored as mutations. No sequences had insertions. Three sequences had deletions; the deleted bases were not counted in the analysis. The Kabat CDR and FR locations as depicted in the V BASE Sequence Directory (Tomlinson et al.) were used. Only the nonproductively rearranged sequences were included in this analysis.

All data were organized as 2 × 2 contingency tables and analyzed using the Yates-corrected chi-square.

Results and Discussion

Among all the sequences examined, 115 have been found to be nonproductively rearranged; these sequences may be obtained by e-mail request to kepler@stat.ncsu.edu. Nonfunctionally rearranged V regions are not expressed, thus the mutation patterns are unselected and therefore directly reflect inherent mutation biases.

We found significant enhancement of the mutation frequency in CDR over that in FR at all three Ig loci (see Table 1). The size of the observed enhancement, 58%, is larger than but consistent with that predicted by differential codon use (Kepler 1997).

Because the hot and cold serine codons differ from each other by more than one base change, the possibility exists that CDR–FR mutability differences are "frozen accidents" resulting from conservation of the serine residues and, by extension, their codon classes (Kepler 1997). We removed this confounding effect by reanalyzing the data, ignoring all serine codons, both mutated and nonmutated (see Table 1). For all loci, the CDR remain more mutated when serine codons are ignored; the λ locus remains highly significant, the κ locus becomes borderline, and the heavy-chain locus, with many fewer events, is nonsignificant.

It is a tantalizingly open question whether this type of adaptation has occurred for other genes, but it is now clearly established for immunoglobulins.

Acknowledgments. This work was supported by the Deutsche Forschungsgemeinschaft, by National Science Foundation award MCB-9357637, by a U.S. Department of Education Graduate Assistance in Areas of National Need (GAANN) Fellowship in Biotechnology, and by a Fulbright Grant.

References

Berek C, Kim HJ (1997) B cell activation and development within chronically inflamed synovium in rheumatoid and reactive arthritis. Semin Immunol 9:261–268

Berek C, Milstein C (1987) Mutation drift and repertoire shift in the maturation of the immune response. Immunol Rev 96:23–41

Berek C, Schröder AE (1997) A germinal center-like reaction in the nonlymphoid tissue of the synovial membrane. Ann NY Acad Sci 815:211–217

Berek C, Berger A, Apel M (1991) Maturation of the immune response in germinal centers. Cell 67:1121–1129

Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS (1993) Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: Clustering, polarity, and specific hot spots. Proc Natl Acad Sci USA 90:2385–2388

Dörner T, Brezinschek HP, Brezinschek RI, Foster SJ, Domiati-Saad R, Lipsky PE (1997) Analysis of the frequency and pattern of somatic mutations within nonproductively rearranged human variable heavy chain genes. J Immunol 158:2779–2789

Jacob J, Kelsoe G, Rajewsky K, Weiss U (1991) Intraclonal generation of antibody mutants in germinal centres. Nature 354:389–392

Kepler TB (1997) Codon bias and plasticity in immunoglobulins. Mol Biol Evol 14:637–643

Lebecque SG and Gearhart PJ (1990) Boundaries of somatic mutation

in rearranged immunoglobulin genes: 5′ boundary is near the promoter, and 3′ boundary is ~1 kb from V(D)J gene. J Exp Med 172:1717–1727

Lefranc MP (1997) The international ImMunoGeneTics database http://imgt.cnusc.fr:8104 (Coordinator: Marie-Paule Lefranc, Montpellier, France; lefranc@ligm.crbm.cnrs-mop.fr) [For reference see Nucleic Acids Res 25:206–221 (1997)]

MacLennan ICM (1994) Germinal Centers. Annu Rev Immunol 12: 117–139

McKean D, Hüppi K, Bell M, Staudt L, Gerhard W, Weigert M (1984) Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. Proc Natl Acad Sci USA 81:3180–3184

MacLennan ICM (1994) Germinal Centers. Annu Rev Immunol 12: 117–139

Motoyama N, Okada H, Azuma T (1991) Somatic mutation in constant regions of mouse lambda 1 light chains. Proc Natl Acad Sci USA 88:7933–7973

Paul WE (1993) Fundamental immunology. Raven Press, New York

Reynaud CA, Garcia C, Hein WR, Weill JC (1995) Hypermutation generating the sheep immunoglobulin repertaoire is an antigen-independent process. Cell 80:115–125

Schröder AE, Greiner A, Seyfert C, Berek C (1996) Differentiation of B cells in the nonlymphoid tissue of the synovial membrane of patients with rheumatoid arthritis. Proc Natl Acad Sci USA 93: 221–225

Shlomchik MJ, Litwin S, Weigert M (1989) The influence of somatic mutation on clonal expansion. Prog Immunol 7:415–423

Smith DS, Creadon G, Jena PK, Portanova JP, Kotzin BL, Wysocki LJ (1996) Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. J Immunol 156:2642–2652

Tomlinson IM, Williams SC, Ignatovich O, Corbett SJ, Winter GV BASE Sequence Directory http://www.mrc-cpe.cam.ac.uk/imt-doc/vbase-home-page.html. MRC Centre for Protein Engineering, Cambridge, UK

Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302:575–581

Wagner SJ, Milstein C, Neubergen MS (1995) Codon bias targets mutation. Nature 376:732