Brief Definitive Report

A CONSERVED HUMAN GERMLINE Vκ GENE DIRECTLY ENCODES RHEUMATOID FACTOR LIGHT CHAINS

By VICTOR RADOUX, POJEN P. CHEN, JOSEPH A. SORGE, AND DENNIS A. CARSON

From the Department of Basic and Clinical Research, Scripps Clinic and Research Foundation, La Jolla, California 92037

Rheumatoid factors (RFs) are anti-IgG autoantibodies that may contribute to immune complex–mediated tissue damage in human autoimmune diseases such as rheumatoid arthritis and mixed cryoglobulinemia (1). However, normal humans and mice also synthesize IgM RFs during secondary immune responses. Before one can understand the meaning of sustained RF synthesis in autoimmune states, the genetic basis of RF production in normal populations must be determined.

A group of monoclonal IgM RFs from patients with mixed cryoglobulinemia have been shown to express crossreactive idiotypes (2). The κ light chains of these RFs have homologous variable regions, as determined by immunoblotting with antibodies against hypervariable region peptides, and by amino acid sequence analysis (3, 4). This data suggested that the RF variable regions were encoded by a single κ variable (Vκ) region germline gene, which is highly conserved in the human population.

To verify the germline origin of RF light chains, it was necessary to sequence the intact Vκ RF gene. The present report describes the cloning from normal human placental DNA of a full-length unrearranged gene that encodes the exact light chain variable region found on four separate IgM RFs. These results show that the light chain variable regions of some human RF autoantibodies can be generated without somatic mutation, and that genes capable of encoding RF light chains are present in normal people.

Materials and Methods

Genomic DNA was prepared from a normal human placenta and digested with the following restriction enzymes: Eco RI, Bam HI, Hind III, Kpn I, and Pvu II (New England Biolabs, Boston, MA), according to the manufacturer's directions. Southern blotting was done by standard methods (5). Hybridization was done in 3× SSC, 50% formamide at 42°C, and washes were carried out with 0.1× SSC, 0.1% SDS at 65°C (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate). The probe was a ^{32}P-labelled 865 bp Sau 3AI fragment from a human VκIII germline gene, designated Humkv305/1, and containing the entire coding region, 432 bp upstream and 145 bp downstream. It was isolated from a fetal liver genomic library screened with a human VκIII pseudogene, as described (6).

This work was supported in part by grants AM-25443, AM-35218, CA-36448, and RR-00833 from the National Institutes of Health, Bethesda, MD. V. Radoux is a Fellow of The Canadian Arthritis Society. P. P. Chen is an Investigator of the Arthritis Foundation. This is publication number 4529BCR from the Research Institute of Scripps Clinic.

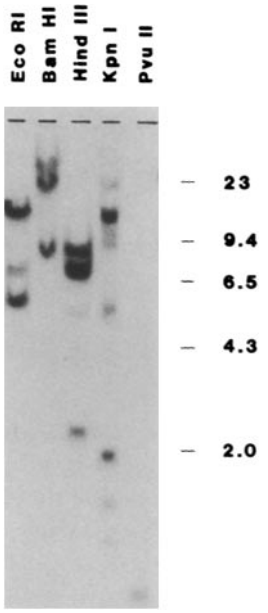

FIGURE 1. Blot hybridization patterns of human germline genomic DNA probed with Humkv305/1. ~10 μg placental DNA was digested with the indicated enzymes. A limited number of hybridizing bands are seen in each case, indicating that few germline genes are closely homologous to Humkv305/1. Size markers (in kb, at right) are from a Hind III digest of λ phage.

Eco RI–digested placental DNA fragments were eluted from the three regions of an agarose gel that corresponded to the Humkv305/1 hybridizing bands, and were used to construct three libraries. The DNA fragments were inserted into the Eco RI sites of the phages λgt10 or EMBL4 (7, 8).Recombinant DNA was packaged into λ phage heads using Gigapack (Stratagene, San Diego, CA). *Escherichia coli* C600 and P2392, respectively, served as hosts. The recombinants were screened with the ^{32}P-labelled Humkv305/1.

The phage clones that hybridized to the probe were isolated and digested with Sau 3AI. Relevant fragments were eluted from low-melting-point agarose gels, subcloned into M13mp8 (9), and sequenced by the dideoxynucleotide chain-termination method (10). Appropriate deletion clones were constructed to complete the sequencing. Computer programs of the University of Wisconsin Genetics Computer Group (UWGCG) were used to analyze sequence data and to prepare figures.

Results

Identification and Isolation of the Vκ RF Gene. To identify germline genes related to the proposed Vκ RF sequence, genomic DNA from a normal human placenta was probed with Humkv305/1, an 865 bp fragment of a human VκIII germline gene. The deduced amino acid sequence of Humkv305/1 shares 92 of 96 amino acids with the putative Vκ RF prototype sequence. When total genomic DNA was digested with a panel of restriction enzymes and then probed with Humkv305/1 under stringent conditions, few bands were identified (Fig. 1). The Eco RI digest revealed three hybridizing bands with approximate sizes of 15, 7.5, and 4.5 kb. DNA was isolated from these three regions and was used to construct libraries in λgt10 or EMBL4, depending upon the fragment size. ~2 × 10^5 recombinant phages from each library were screened with the

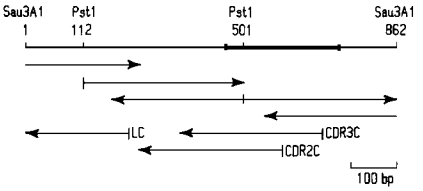

FIGURE 2. Sequencing strategy of the Vκ RF gene. Subclones containing the 862 bp Sau 3A1 fragment, the 389 bp Pst 1–Pst 1 fragment and 361 bp Pst 1–Sau 3A1 fragment were generated. Sequencing primers used included a universal sequencing primer (GTAAAACGACGGCCAGT) and three additional primers, TACCATACTGCTGACAGTAATAC (*CDR3C*), CAGTGGCCCTGCTGGATGCACCA (*CDR2C*), and CCTGGGAGCCAGAGTAGCAGGAG (*LC*), corresponding to pertinent regions of the complementary strand of Humkv305/1 or Humkv321 (6). The thickened segment represents the coding region.

Humkv305/1 probe. The number of clones isolated from each library were respectively four (15 kb), two (7.5 kb), and three (4.5 kb). These clones were subjected to restriction enzyme analysis to confirm their size and relatedness to each other. Then all unique clones were subcloned and sequenced. The Vκ RF gene came from the 4.5 kb fragment cloned into λgt10.

Characterization of the Vκ RF Gene. A Sau 3AI fragment containing the coding region of the Vκ RF was subcloned into M13mp8. Both strands were sequenced. Fig. 2 summarizes the sequencing strategy. The nucleotide sequence of the Sau 3AI fragment of the Vκ RF germline gene is shown in Fig. 3. The two first bases (GA) have been omitted for clarity. The 5′ flanking region contains the pentadecamer and decamer sequences that are important for transcription of Ig genes (11). Between the two sequences is the octamer CCCGCCCC, which represents a potential binding site for the Sp1 transcriptional regulatory protein (12). The heptamer and nonamer sequences for V-J rearrangement are also present. The deduced amino acid sequence is indicated in Fig. 3 by the three-letter code, and is identical to the protein sequence from four separate IgM-RF light chain variable regions (4). Thus, the Vκ RF is a functional gene. The close relation to Humkv305/1 is confirmed by an overall sequence homology of 98%.

Discussion

The sequence determination of the entire Vκ RF gene proves that the light chain variable region of human anti-IgG autoantibodies is encoded by a germline gene without any somatic mutation. The Vκ RF gene was isolated from a normal human placenta; a fragment of this gene was previously isolated from a fetal liver library (6). Thus, the Vκ RF gene is not confined to patients with IgM RF cryoglobulins, but is also present in normal people.

When Southern blots of Eco RI–digested DNA from multiple individuals were probed with either the Humkv305/1 or Vκ RF genes, the same three bands hybridized. However, it is still possible that the exact structure of the Vκ RF gene varies in outbred human populations. The variants could represent: (*a*) alleles at the Vκ RF locus, (*b*) related members of the same gene family as the Vκ RF and Humkv305, (*c*) somatic mutants. To distinguish between these possibilities will require the cloning and sequencing of the Vκ RF hybridizing DNA bands from both normal subjects and patients with autoimmune disease.

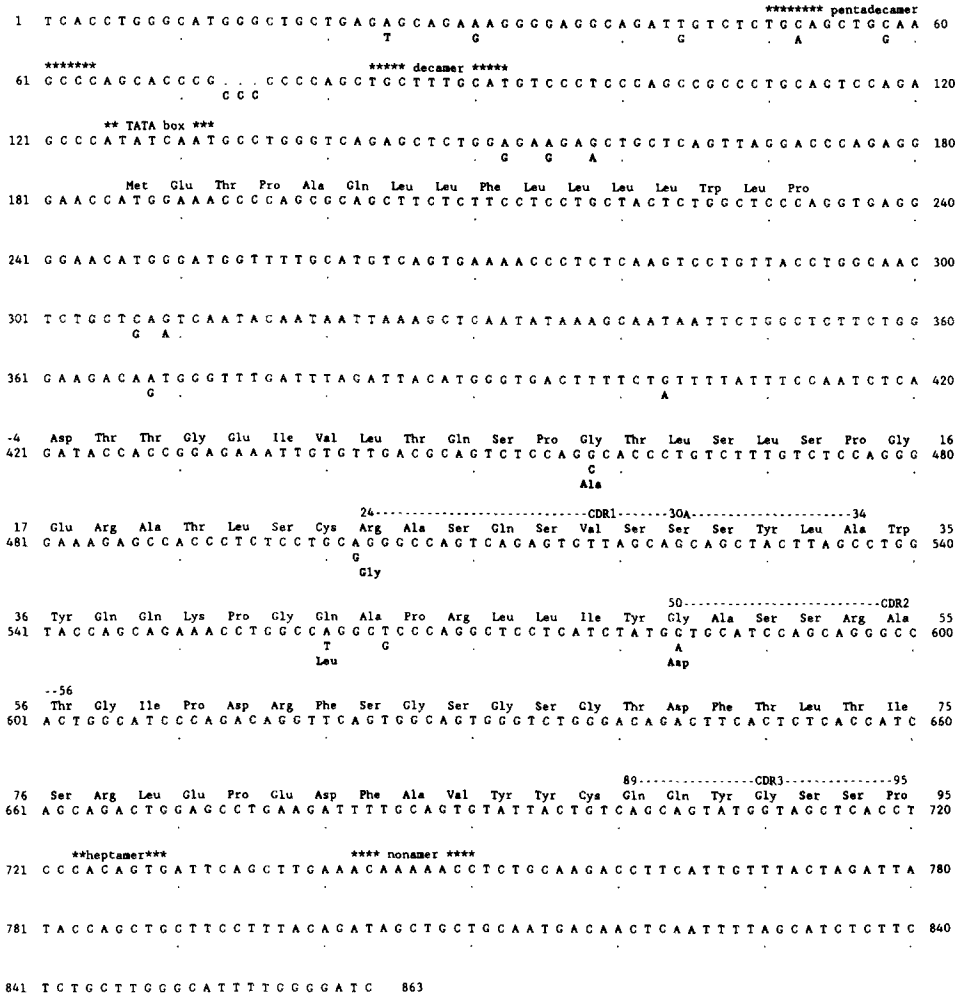

FIGURE 3. Sequence of Vκ RF compared with Humkv305/1. Nucleotides and the deduced amino acids of the Vκ RF gene appear above, and differences with Humkv305/1 are shown below. Dots inserted at positions 73–75 were introduced to maximize homology with Humkv305/1. The pentadecamer, decamer, TATA box, leader sequence, heptamer, and nonamer are indicated. Complementarity-determining regions (CDRs) are indicated, and the extra codon for the CDR1 is placed at position 30A.

The results should reveal whether or not polymorphisms in the coding or flanking regions of the Vκ RF gene are a risk factor for the development of autoimmunity.

Information concerning the V_H genes that encode human IgM RF autoantibodies is still meager. The amino acid sequences of two RF heavy chain variable regions belonging to the Wa idiotypic group were found to be different (13). Similarly, antibodies against hypervariable region peptides corresponding to some RF heavy chain sequences failed to react with most other RF paraproteins (14). These results can be interpreted to suggest that the RF V_H regions contain somatic mutations. However, another possible explanation is that RF V_H regions are encoded by several V_H genes in the germline immune repertoire. Such an

interpretation is consistent with the constitutive presence of abundant RF precursor cells in human umbilical cord blood (1).

The potential physiologic functions of IgM RFs have been reviewed recently (1). The anti-IgG autoantibodies may amplify immune responses to infectious agents by crosslinking IgG aligned on a cell surface and by enhancing complement fixation. The RFs may also promote the clearance from the circulation of aggregated and denatured IgG.

The four monoclonal RFs containing the Vκ RF sequence were isolated from patients with mixed cryoglobulinemia (4). It may seem paradoxical that the anti-IgG autoantibodies that produce vascular damage in these patients use the same conserved Vκ RF gene that presumably encodes the physiologic RFs in healthy people. However, the important point is that individuals with cryoglobulinemia produce massive quantities of RF. The recurrent use of the Vκ RF gene in mixed cryoglobulinemia may relate to its structure or location. The most J_H-proximal V_H gene segments are preferentially rearranged in Abelson virus–transformed murine B cell leukemias (15). It is conceivable that a similar phenomenon could influence the frequency of rearrangements, or the expression, of the human Vκ RF gene locus.

Recent experiments in mice have demonstrated that both germline and somatically mutated Ig variable region genes may encode anti-DNA antibodies (16, 17). The same situation probably applies to human IgM RFs. Thus, a major fraction of the RFs in patients with Sjogren's syndrome express Vκ RF–associated crossreactive idiotypic antigens (18). In contrast, only a minor proportion of the IgM RFs in patients with rheumatoid arthritis apparently use the Vκ RF gene in its unmutated form. Future experiments will need to assess more precisely the relative roles of variable region gene polymorphisms, and antigen-driven diversification mechanisms, in the genesis of both physiologic and pathogenic human autoantibodies.

Summary

The full-length gene that encodes the light chain variable regions of an idiotypically related group of human IgMκ rheumatoid factors (RFs) has been cloned and sequenced. The deduced amino acid sequence is identical to four separate RF proteins. These results prove that genes capable of encoding human anti-IgG autoantibody light chains without any somatic mutation are present in the κ gene repertoire of normal people.

We thank Emily Y. Chen for making synthetic oligomers, Heidi Short and Robert Shrantz for technical assistance, and the Basic and Clinical Research Word Processing Center for their help in the preparation of this manuscript.

Received for publication 2 September 1986 and in revised form 29 September 1986.

References

1. Carson, D. A., P. P. Chen, R. I. Fox, T. J. Kipps, F. Jirik, R. D. Goldfien, G. Silverman, V. Radoux, and S. Fong. 1986. Rheumatoid factors and immune networks. *Annu.Rev. Immunol.* In press.
2. Kunkel, H. G., V. Agnello, F. G. Joslin, R. J. Winchester and J. D. Capra. 1973.

Cross-idiotypic specificity among monoclonal IgM proteins with anti-gammaglobulin activity. *J. Exp. Med.* 137:331.

3. Chen, P. P., F. Goni, S. Fong, F. Jirik, J. H. Vaughan, B. Frangione, and D. A. Carson. 1985. The majority of human monoclonal IgM rheumatoid factors express a primary structure-dependent cross-reactive idiotype. *J. Immunol.* 134:3281.
4. Goni, F., P. P. Chen, B. Pons-Estel, D. A. Carson, and B. Frangione. 1985. Sequence similarities and cross-idiotypic specificity of L chains among human monoclonal IgM-K with anti-gammaglobulin activity. *J. Immunol.* 135:4073.
5. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 382–389.
6. Chen, P.P., K. Albrandt, N. K. Orida, V. Radoux, E. Y. Chen, R. Schrantz, F.-T. Liu, and D. A. Carson. 1986. Genetic basis for the cross-reactive idiotypes on the light chains of human IgM anti-IgG autoantibodies. *Proc. Natl. Acad. Sci. USA.* In press.
7. Huynh, T. V., R. A. Young, and R. W. Davis. 1985. Constructing and screening cDNA libraries in lambda-gt10 and lambda-gt11. DNA Cloning: A Practical Approach, vol. I. D. M. Glover, editor. IRL Press, Oxford, United Kingdom. 49–78.
8. Frischauf, A. M., H. Lehrach, A. Poustka and N. Murray. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170:827.
9. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20.
10. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
11. Falkner, F. G., and H. G. Zachau. 1984. Correct transcription of an immunoglobulin K gene requires an upstream fragment containing conserved sequence elements. *Nature (Lond.).* 310:71.
12. Kadonaga, J. T., K. A. Jones and R. Tjian. 1986. Promoter-specific activation of RNA polymerase II transcription by SpI. *TIBS (Trends. Biochem. Sci.).* 11:20.
13. Andrews, D. W., and J. D. Capra. 1981. Complete amino acid sequence of variable domains from two monoclonal human anti-gamma globulins of the Wa cross-idiotypic group: Suggestion that the J segments are involved in the structural correlate of the idiotype. *Proc. Natl. Acad. Sci. USA,* 78:3799.
14. Chen, P. P., F. Goni, R. A. Houghten, S. Fong, R. Goldfien, J. H. Vaughan, B. Frangione and D. A. Carson. 1985. Characterization of human rheumatoid factors with seven antiidiotypes induced by synthetic hypervariable region peptides. *J. Exp. Med.* 162:487.
15. Yancopoulos, G. D., S. V. Desiderio, M. Paskind, J. F. Kearney, D. Baltimore, and F. W. Alt. 1984. Preferential utilization of the most J_H-proximal V_H gene segments in pre-B-cell lines. *Nature (Lond.).* 311:727.
16. Diamond, B., and M. D. Scharff. 1984. Somatic mutation of the T15 heavy chain gives rise to an antibody with autoantibody specificity. *Proc. Natl. Acad. Sci. USA.* 81:5841.
17. Naparstek, Y., J. Andre-Schwartz, T. Manser, L. J. Wysocki, L. Breitman, B. D. Stollar, M. Gefter, and R. S. Schwartz. 1986. A single germline V_H gene segment of normal A/J mice encodes autoantibodies characteristic of systemic lupus erythematosus. *J. Exp. Med.* 164:614.
18. Fong, S., P. P. Chen, T. A. Gilbertson, J. R. Weber, R. I. Fox, and D. A. Carson. 1986. Expression of three cross-reactive idiotypes on rheumatoid factor autoantibodies from patients with autoimmune diseases and seropositive adults. *J. Immunol.* 137:122.