Genetic basis for the cross-reactive idiotypes on the light chains of human IgM anti-IgG autoantibodies

(anti-peptide antibody/human immunoglobulin genes)

Pojen P. Chen*, Keith Albrandt†, Norman K. Orida†, Victor Radoux*, Emily Y. Chen‡, Robert Schrantz*, Fu-Tong Liu†, and Dennis A. Carson*

Departments of *Basic and Clinical Research and †Molecular Biology, Scripps Clinic and Research Foundation, 10666 North Torrey Pines Road, La Jolla, CA 92037; and ‡Division of Molecular Biology, Medical Biology Institute, 11077 North Torrey Pines Road, La Jolla, CA 92037

Communicated by Ernest Beutler, July 11, 1986

ABSTRACT The role of immunoglobulin structural genes in the generation of autoantibodies in humans has not been elucidated. Human monoclonal IgM anti-IgG autoantibodies (rheumatoid factors, RFs) from unrelated people often share idiotypic antigens. Antibodies against synthetic peptides have localized two of the shared idiotypic determinants to the second and third complementarity-determining regions of the κ light chain. The reported sequences of several human RF light chains are remarkably homologous in these regions. Animal studies have shown that some shared idiotypic antigens represent serological markers for immunoglobulin variable (V)-region genes. Therefore, we hypothesized that human RF light chains derived from a single germ-line gene, designated Vκ(RF), or from a small family of very closely related genes. In the present experiments, we have isolated and sequenced two human Vκ germ-line genes that encode κ light chains, which are identical or closely related to the light chains of human RF. The data indicate that the shared idiotypic antigens on RF are phenotypic markers for a κ V-region gene that is highly conserved in the human population. The results also imply that the light chains of IgM anti-IgG autoantibodies can be encoded by germ-line genes without any somatic mutation.

MATERIALS AND METHODS

Genomic Library, Probes, and Screening. The human genomic library was kindly provided by T. Maniatis (Harvard University), and was constructed by Hae III/AIu I partial digestion of fetal liver DNA, followed by ligation with EcoRI linkers and insertion into λ Charon 4A (17). A total of 10⁶ clones from a secondary amplification was first screened with the human Vκ cDNA subclone NG9/3 (18, 19). The NG9/3 consists of a 243-base-pair (bp) fragment corresponding to amino acids 4 to 84 in a κ IIIb light chain and was obtained from D. L. Bentley (Fred Hutchinson Cancer Research Institute, Seattle, WA). One recombinant phage that hybridized strongly with NG9/3 was found to contain a pseudogene belonging to the κ III subgroup, and the gene was thus designated Humkv301. An Sau3AI fragment of 895 bp, designated Humkv301/1 and containing nucleotides from position −96 to +799, was subcloned into M13mp8 and used to rescreen the entire original library. Probes were labeled to a specific activity of ≈10⁸ cpm/μg by nick-translation using [32P]dCTP (19).

DNA Sequencing. Initially, appropriate Sau3AI fragments were enriched from low-melting-point agarose gels and were ligated into M13mp8 according to standard methods (20, 21). The clones containing the coding region sequences were identified by hybridization with NG9/3 and sequenced by the dideoxynucleotide chain-termination method of Sanger et al. (20). Subsequently, additional clones containing appropriate restriction fragments were constructed and used for sequencing both DNA strands. The computer programs of the University of Wisconsin Genetics Computer Group were used to assemble, edit, and analyze all sequence data and to prepare figures.

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.

Abbreviations: bp, base pair(s); RF, rheumatoid factor; V region, variable region of immunoglobulin.
RESULTS

Identification and Isolation of Humkv305 and Humkv321. The amino acid sequence corresponding to the cDNA subclone NG9/3 differs from the prototype sequence of human RF light chains by three amino acids in the V\(_s\) region (amino acid residues 1-95) (Fig. 1). All three different amino acids can be explained by single base changes. The NG9/3 cDNA was isolated from a light chain secreting human-mouse heterohybridoma with no known antibody specificity (23). To ensure that the NG9/3 cDNA would identify all relevant germ-line \(V_s\) genes, library screening was done at low stringency in 2\(\times\) SSC (1\(\times\) SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0) at 65°C, followed by washing in 1\(\times\) SSC at 65°C.

Among 10\(^6\) recombinant phage screened by this method, 98 positive clones were identified. Subsequently, one recombination phage with a strong hybridization signal was subcloned into M13mp8 and sequenced. This clone, designated Humkv301, contained a pseudogene belonging to the \(s\) III subgroup (unpublished data).

Since immunoglobulin V\(_{\alpha}\)-region genes are generally more heterogeneous in the flanking regions than in the coding regions, an Sau3AI fragment of Humkv301 (from position -96 to +799, including both the flanking regions and the whole coding region) was used to reprobe the original set of library screening blots. A total of 21 clones were identified among the 98 clones that hybridized with NG9/3. The 21 clones were classified into seven groups on the basis of Southern blot analyses with the NG9/3 probe. One clone from each group was subcloned into M13 and sequenced. Clones Humkv305 and Humkv321, from two groups, were most relevant to the postulated germ-line sequence of the monoclonal RF light chain (Fig. 1).

Characterization of Humkv305 and Humkv321. The Sau3AI digest of both Humkv305 and Humkv321 clones yielded a band of \(\approx 1\) kilobase. These fragments were subcloned into M13mp8 for sequencing. Initially, clones with fragments inserted in both orientations were sequenced with the universal sequencing primer (GTAAACACGGCAGCT; FR3C) corresponding to the complementary strand of the third framework region of NG9/3. To sequence both strands completely, deletion clones of the Humkv305 were generated using either Pst I or HincII. In addition, clones containing the EcoRI fragment of Humkv321 in both orientations were prepared. These clones were sequenced with the universal sequencing primer and an additional primer (CTCATCCATG; FR2S) corresponding to the sense strand of the second framework region of the Humkv301 (Fig. 2).

As illustrated in Fig. 1, the 54 amino acids encoded by Humkv321 correspond exactly to the postulated sequence of the idiotype-positive monoclonal RF light chains from position 42 to 95. Fig. 3 shows the sequence of a 394-base fragment, which contains 306 bp of the Humkv321 gene. The first 85 nucleotides match completely to nucleotides from positions 4302-4386 of the Escherichia coli lacZ gene (24), which was inserted into the cloning vector for indicator function (25). Thus, during generation of the library the Humkv321 gene was cut at an internal Hae III site (GGCC) (see the corresponding site in Humkv305 in Fig. 4), and cloned into the EcoRI site of the phage vector.

Clone Humkv305 contains the appropriate consensus sequences for transcription and variable-joining region (V-J) rearrangement (26, 22), and shares 92 of 96 amino acids with the proposed sequence of the idiotype-positive monoclonal RF light chains (Figs. 1 and 4). It differs from the prototype at positions 9, 24, 42, and 50 (designated in Fig. 4 by lower-case letters). Each amino acid difference can be explained by a single base change. Importantly, two previously reported germ-line \(s\) III light chain genes, Humkv3g and Humkv3h (22), differ from this sequence by 10 and 13 amino acids, respectively (Fig. 1).

Fig. 1. Comparison of amino acid sequences (single-letter code) in the \(V_s\) gene-encoded regions (positions 1-95). The prototype sequence is from Chen et al. (15). The cDNA NG9 was cloned and sequenced by Bentley (18). The threonine at position 52 was changed to serine based on the corrected nucleotide sequence (GCA TCC, instead of GCT ACC; positions 154-159; P.P.C and D.A.C., unpublished data). The Humkv3g and Humkv3h are renamed from Vg and Vh, which were cloned and sequenced by Pech and Zachau (22).

Fig. 2. DNA sequencing strategy. For Humkv305, the clones containing the 865-bp Sau3AI fragment (F), the 750-bp Pst I/Sau3AI F, the 445-bp Sau3AI/HincII F, the 420-bp HincII/Sau3AI F, and the 115-bp Sau3AI/Pst I F were used in sequencing. For Humkv321, the clones containing the 1000-bp Sau3AI F (starting at position 18,783 of the \(\lambda\) phase) and the 314-bp EcoRI/Sau3AI F were used. Except where indicated (with the FR2S primer and the FR3C primer; see Results), all clones were sequenced with the universal sequencing primer.
Fig. 3. The partial genomic structure of the *Humkv321* gene. Nucleotides of the *Humkv321* are upper case; nucleotides of the phase vector and the linker are lower case. The first 85 nucleotides match completely to positions 4386–4302 of the *E. coli lac*Z gene (24), which was inserted into λ Charon 4A for index function (25). Amino acids are designated by the single-letter code.

**DISCUSSION**

The sharing of idiotypic antigens by RFs from unrelated subjects prompted us to hypothesize that genes related to the primary sequence of the autoantibodies were widely distributed in the human genome (15). The current isolation of two human germ-line *V*~e~ genes that encode sequences related to those on the light chains of monoclonal RFs provides firm support for this supposition. The translated sequence of one cloned gene (*Humkv321*) is identical to the proposed RF.

**FIG. 4.** The genomic structure of the *Humkv305* gene. Nucleotides are numbered according to the first nucleotide (line 1) or to the translated amino acid sequence (line 2). The 15-mer, 10-mer, TATA box, and 7-mer are the respective consensus sequences for gene rearrangement and transcription and are marked according to Pech and Zachau (22). Dots designate insertions, which are introduced to maximize homology among four human *κ* germ-line genes. The sites recognized by enzymes used in generating deletion clones are indicated. In addition, the *Hae* III site corresponding to amino acid position 41, where Humkv321 was cut during the generation of the library, is indicated. Amino acids are designated by the single-letter code.
light-chain prototype sequence from position 42 to 95. This part of the gene includes two cross-reactive idiotypic determinants that were defined by antibodies against synthetic peptides corresponding to the second and third complementarity-determining regions (15, 16, 27).

Fig. 5 summarizes the reported amino acid sequences of nine idiotype-positive human RF light chains, as determined by Frangione, Capra, and their colleagues (11–13, 28, 29). Notably, four of the light chains are identical to the proposed germ-line $V_\kappa$(RF) sequence. Taken together with the present data, the results strongly suggest that a $V$-region germ-line gene can encode human RF light chains without any somatic mutation.

It seems likely that the isolated RF $V_\kappa$ gene, or closely related genes, are present in nearly all humans. Thus, every serum that we have tested contains some light chains that share idiotypic antigens with monoclonal RFs (30–32). Furthermore, we recently generated multiple human B-cell hybridomas that share idiotypic antigens with RFs. All the idiototype-positive hybridomas contain rearranged $V_\kappa$ genes that are extremely homologous to the germ-line RF prototype sequence (R. D. Goldfien, S. Fong, J. G. Heitmann, V.R., P.P.C., and D.A.C., unpublished data).

In addition to the germ-line $V_\kappa$ genes described here, we have recently cloned a rearranged $V_\kappa$ gene from a RF-secreting human B-cell neoplasm (33). At least two monoclonal RFs from unrelated individuals have sequences closely related to this gene in the second and third complementarity-determining regions (10). These data suggest that a second conserved germ-line $V_\kappa$ gene may also be utilized for RF light-chain synthesis.

The assembly of a functional antibody molecule requires the interaction of light and heavy chains that are encoded by genes on separate chromosomes. Limited experiments with anti-peptide antibodies have failed to reveal homologous sequences in the complementarity-determining regions of RF heavy chains (16). Other studies have shown that some antibodies with idiotype-positive light chains do not bind human IgG (13, 30). Collectively, these results suggest that only a selected group of heavy-chain genes can be used for RF synthesis.

Hybridization studies indicate that the human germ line contains only 25–50 $V_\kappa$ genes (18, 26, 34, 35). Strong selective pressures must have operated to conserve the basic RF $V_\kappa$ sequence during the evolution and dispersion of the human species (36, 37). Possibly, the $V_\kappa$(RF)-encoded light chain, in combination with an appropriate heavy chain, generates an antibody against an important environmental pathogen. However, recent experiments indicate that normal humans and animals regularly synthesize IgM RFs during secondary immune responses (37–39). By cross-linking immunoglobulin constant regions, the RFs magnify the complement-fixing effects of IgG antibodies against cell-surface antigens. The ability of RFs to promote the removal of opsonized bacteria and parasites may be the reason why this autoantibody has been maintained during evolution (40).

Jerne postulated that the antibody repertoire is regulated normally by idiotype-antiidiotype interactions (41, 42). Some RFs have been shown to recognize antigenic determinants in both the Fab and Fc fragments of the IgG molecule (43, 44). Since both the variable and constant regions of immunoglobulins evolved from a common ancestor domain, they may share antigens in common. It is therefore conceivable that some idiotypes derive from the germ-line-encoded RFs through somatic mutation. This hypothesis, if substantiated, is another explanation for the conservation of the $V_\kappa$(RF) genes in outbred human populations.

RFs that express the germ-line-encoded idiotypic antigens are especially common among Waldenstrom macroglobulins. The same idiotypic determinants are infrequently found

\[
\begin{array}{cccc}
2 & 3 & 3 \\
4 & 0 & 4 \\
A \\
\end{array}
\]

$V_\kappa$(RF) sequence (15) is shown on the top. Dashed lines represent amino acids that are identical with the proposed sequence. All nine light chains share two idiotypic antigens in the second and third complementarity-determining regions (15, 16, 28, 29).

Fig. 5. Amino acid sequences (single-letter code) of nine idiotype-positive human RF light chains (11, 13, 28, 29). The proposed $V_\kappa$(RF) sequence (15) is shown on the top. Dashed lines represent amino acids that are identical with the proposed sequence. All nine light chains share two idiotypic antigens in the second and third complementarity-determining regions (15, 16, 28, 29).
among IgG myelomas or in serum IgG. One explanation for this phenomenon is that neoplastic transformation in patients with Waldenstrom macroglobulinemia, and perhaps in many patients with chronic lymphatic leukemia, occurs at an early stage of B-cell development before somatic mutations have accumulated. Considering the relatively small number of germ-line V\(_\kappa\) genes in humans, it is not surprising that malignancies of immature B cells express recurrent idio-
types.

In summary, these experiments show that the cross-
reactive idiootypes on human RFs are phenotypic markers of a germ-line gene that is widespread in the normal human population. Furthermore, IgM anti-IgG autoantibody light chains can be encoded by germ-line genes without any somatic mutation.

The authors wish to thank Drs. T. Maniatis and D. L. Bentley for providing the human genomic library and the cDNA NG9/3 subclone, Arpana Ewing for assisting us in running the computer programs, and Drs. Joe A. Sorge and Christine A. Gritzmacher for helpful advice. We acknowledge the excellent assistance of the BCR Word Processing Center in the preparation of this manuscript. Funding for this research was supported in part by Grants AM 35218, AM 25442, AI 19747, and RR 00833 from the National Institutes of Health. This is publication 4194BCR from the Research Institute of Scripps Clinic and Research Foundation 87 from the Medical Biology Institute. P.P.C. is an Investigator of the Arthritis Foundation. F.-T.L. is a Scholar of the Leukemia Society of America. V.R. is a Fellow of the Canadian Arthritis Society.

   no nephrol. 1, 133–169.
   177–179.
   569–607.
    Pasteur 127C, 261–271.
    USA 78, 3799–3803.
12. Ledford, D. K., Gofi, F., Pizzolato, M., Franklin, E. C.,
    1322–1325.
14. Chen, P. P., Houghten, R. A., Fong, S., Rhodes, G. H.,
    Gilbertson, T. A., Vaughan, J. H., Lerner, R. A. & Carson,
15. Chen, P. P., Gofi, F., Fong, S., Jirik, F., Vaughan, J. H.,
    3281–3285.
16. Chen, P. P., Gofi, F., Houghten, R. A., Fong, S., Goldfien,
17. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. &
    Cloning: A Laboratory Manual (Cold Spring Harbor Labora-
    tory, Cold Spring Harbor, NY).
    Acad. Sci. USA 74, 5463–5467.
    9229–9236.
    EMBO J. 2, 593–597.
25. Blattner, F. R., Williams, B. G., Blechier, A. E., Denniston-
    Thompson, K., Faber, H. E., Furlong, L., Grunwald, D. J.,
27. Chen, P. P., Fong, S., Normansell, H., Houghten, R. A.,
    Med. 159, 1502–1511.
29. Newkirk, M., Chen, P. P., Carson, D. A., Posnett, B.
30. Fong, S., Chen, P. P., Gilbertson, T. A., Fox, R. I., Vaughan,
31. Fong, S., Chen, P. P., Gilbertson, T. A., Weber, J. R., Fox,
32. Chen, P. P., Fong, S., Gofi, F., Houghten, R. A., Frangione,
    B., Liu, F. & Carson, D. A. (1986) in Idiotypes and Diseases,
    eds. Zanetti, M., Bona, C. A. & Celada, F. (Karger, Basel,
    Switzerland), in press.
33. Jirik, F. R., Sorge, J., Fong, S., Heitzmann, J. G., Chen,
36. Carson, D. A., Pasquali, J.-L., Tsoukas, C. D., Fong, S.,
    Slovin, S. F., Lawrence, S. K., Slaughter, L. & Vaughan,
    Gerontology 31, 236–250.
    895–899.
    529–545.
    373–387.
    nobiol. 125, 1530–1535.
    Springer Semin. Immunopathol. 6, 33–49.