GENETIC ELEMENTS USED FOR A MURINE LUPUS ANTIDNA AUTOANTIBODY ARE CLOSELY RELATED TO THOSE FOR ANTIBODIES TO EXOGENOUS ANTIGENS

BY REINHARD KOFLER, DANIEL J. NOONAN, DAVID E. LEVY, MICHAEL C. WILSON, NIELS P. H. MØLLER, FRANK J. DIXON, AND ARGYRIOS N. THEOFILOPOULOS

From the Departments of Immunology and Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, California 92037

Much of our understanding of the generation of antibody diversity stems from recombinant DNA studies on myeloma or hybridoma antibodies with unknown specificity, or specificity for exogenous antigens (reviewed in 1). Hence, the antibody response to a number of defined haptens in different mouse strains has been delineated with respect to the germline elements used and the contribution of somatic events to creation of antibody diversity (2–6). Autoantibodies involved in autoimmune disease pathogenesis have not been well studied at the molecular level, and their genetic origin, inducing agents, and means of circumventing self-tolerance remain largely unknown. To determine their structural genetic elements (variable [V], diversity [D], joining [J], and constant [C] gene segments), we initiated molecular analysis of human and murine hybridoma-derived monoclonal autoantibodies against several self-antigens (7, and unpublished data). Of particular interest was whether autoantibodies use unique Ig genes or derive from the same germline pool as antibodies to exogenous antigens.

We report herein the molecular cloning and complete nucleic acid sequence of mRNAs corresponding to the heavy (H) and light (L) chains of a monoclonal IgMk anti-DNA autoantibody from the MRL/Mp-lpr/lpr (MRL/l) strain, which spontaneously develops a systemic lupus erythematosus (SLE)-like syndrome associated with circulating anti-DNA autoantibodies (8). Comparisons with published antibody sequences indicate that the H chain of this autoantibody originates from a Vμ gene of the 4-hydroxy-3-nitrophenyl (NP) antibody family (2, 9), and its L chain from a V gene of the Vλ1 group, also found in antibodies to exogenous antigens (10, 11). The J segments closely resemble BALB/c Jm3 and Jα sequences, respectively, and the deduced C region amino acid sequences are identical to those found in non–self-directed IgMk immunoglobulins. Only the D
segment, which is frequently subject to extensive somatic alterations (1), could not be classified. Our findings suggest that, with the possible exception of the D segment, this anti-DNA autoantibody and, by extension, perhaps autoantibodies in general, do not originate from autoantibody-specific structural genetic elements, but from the same germline repertoire as antibodies to exogenous antigens.

Materials and Methods

The hybridoma cell line MRL-DNA10 originated from a fusion of 3-mo-old MRL/l mouse spleen cells with the BALB/c nonsecretor myeloma cell line P3-X63-Ag8.653. From this fusion, seven anti-DNA autoantibody-secreting hybridoma lines (1.4% of growing hybrids) were recovered. The details of their production and characterization will be the subject of a separate report. MRL-DNA10 secretes a monoclonal IgM, autoantibody with specificity for single-stranded DNA. Autoantibodies of this specificity occur in the sera of MRL/l mice at high titers (8).

Total cellular RNA was prepared from MRL-DNA10 cells by the isothiocyanate technique (12). Poly(A)+ RNA was purified by affinity chromatography over oligo(dT) columns (13) and copied into cDNA using the Okayama-Berg cloning procedure (14). Escherichia coli (strain MC1061) were transformed with the resulting plasmids by the CaCl2 technique (15), cloned on ampicillin-containing agarose, and transferred to nitrocellulose filters. Replica filters were screened for clones containing H or L chain cDNA by hybridization to 32P-labeled C region probes for murine μ (ChSpa7) (16) and κ (pCRI-k40) (17) chain sequences, respectively, using standard procedures (18). Plasmid DNA from positive colonies was prepared by alkaline lysis (19) and digested with a series of restriction enzymes, and restriction enzyme maps for H and L chain cDNA species were established (Fig. 1). Appropriate fragments were labeled with [32P]deoxynucleotide triphosphates on their 3' end with Klenow fragment of DNA polymerase or their 5' end with polynucleotide kinase, in each case cut with an appropriate second restriction enzyme, and purified on polyacrylamide gels. The eluted fragments were then sequenced by partial chemical degradation (20).

Automated N-terminal amino acid sequence analysis was performed on an updated model 890M sequencer (Beckman Instruments, Inc., Fullerton, CA) and phenylthiohydantoin amino acids were identified by high pressure liquid chromatography (21). The N-terminal L chain sequence was obtained from the unseparated IgM molecule since the amino terminus of the H chains was blocked. For H chain sequencing, affinity-purified MRL-DNA10 protein was reduced with dithiothreitol, alkylated with iodoacetic acid (22), and H and L chains separated by high pressure liquid chromatography. The isolated H chains were deblocked with pyroglutamate aminopeptidase (23) and subjected to automated sequence analysis. The nucleic acid and deduced amino acid sequences were compared with immunoglobulin sequences from the National Biomedical Research Foundation database with the computer program ALIGN (24).
Results

Structural Genetic Elements Creating the MRL-DNA10 H Chain. ~1% of the 4,000 to 5,000 E. coli colonies transformed with MRL-DNA10 cDNA hybridized to the C\textsubscript{\textmu} probe. Restriction enzyme analysis of plasmid DNA from 10 \mu-positive clones suggested the presence of a single IgM H chain mRNA species (data not shown). Three clones contained cDNA inserts of ~2.3 kilobases, which corresponds to a full-length secreted IgM message. The entire insert sequence was determined and identified as a copy of an IgM H chain mRNA, based on the similarity of its 3' portion to a published C\textsubscript{\textmu} region DNA sequence (see below). The correspondence of the respective message with the MRL-DNA10 IgM H chain protein was further supported by matching of the first 12 amino acids, as determined by N-terminal sequencing of the deblocked MRL-DNA10 H chain, with corresponding residues deduced from the cDNA nucleotide sequence. This included a proline in position 7 that is rarely found in H chain sequences (9).

Computer comparison of the sequence encoding the 5' untranslated region, leader peptide, and amino acids 1-94 showed a high degree of similarity with members of a \textit{V\textsubscript{\textmu}} gene family encoding the predominant antibodies to NP in C57BL/6 (Igh\textsuperscript{b}) mice (9) and with members of a corresponding gene family from the Igh\textsuperscript{h} haplotype BALB/c strain (2). The most closely related \textit{V} genes from each of the two Igh haplotypes are depicted in Fig. 2 in comparison with the MRL-DNA10 sequence. Similarity was greatest (389 of 397 basepairs, or 98%) with the C57BL/6 germline pseudogene 6, which is nonfunctional due to a single basepair deletion in codon 22. Since the MRL/I strain carries the Igh\textsuperscript{h} haplotype in the respective locus, the germline gene encoding MRL-DNA10 could be a functional Igh\textsuperscript{b} allele of the C57BL/6 pseudogene 6. In fact, this anti-DNA autoantibody crossreacted with NP and dinitrophenyl in a radioimmunoassay, but did not bind nitroiodophenyi (Morrow and Kofler, unpublished observation), distinguishing it from the heteroclytic NP antibodies of Igh\textsuperscript{b} haplotype (25).

The J\textsubscript{\textmu} segment used by the MRL-DNA10 H chain corresponded to the BALB/c J\textsubscript{\textmu}3 sequence, with the exception of the first nucleotide (GGG instead of TGG in codon 99; Fig. 2). The close similarity to BALB/c (Igh\textsuperscript{b}) was not unexpected, since a BALB/c-identical J\textsubscript{\textmu}3 segment occurred in an antiphosphocholine (anti-PC) antibody protein sequence from the Igh\textsuperscript{b} strain CBA/J (3). This nucleotide difference produced an amino acid substitution (glycine for tryptophan), and may have resulted from somatic events that occur frequently at this position (1). The relatively short putative D segment (codons 95-97; codon 98 was absent) did not show distinct similarity to D segments available for comparison.

The entire C\textsubscript{\textmu} coding region (data not shown) differed from a genomic BALB/c IgM sequence described by Goldberg et al. (26) in only four nucleotides (codons 122-TTC, 143-CTA, 426-AGC, and 525-TTG were, instead, TTT, CTG, AGT, and CTG in the MRL-DNA10 sequence). All differences were silent, i.e., they did not lead to an amino acid change. Another nucleotide difference was found in the 3' untranslated region (at position 1991, T [Ref. 26])

FIGURE 2. Nucleotide sequence and deduced amino acid sequence of the MRL-DNA10 V_{H} region compared with the C57BL/6 NP-V_{H} gene 6 (Igh^{b}) and the BALB/c "NP-equivalent," V_{H} gene 3 (Igh^{a}). Dotted lines signify identity with the MRL-DNA10 nucleotide sequence. Amino acids differing from the deduced V_{H} gene 6 sequence (replacing the deletion in codon 22 by T) are indicated by asterisks. For clarity, differences from the deduced V_{H} gene 3 amino acid sequence are not marked. The putative D and J_{H} gene encoded sequences and corresponding amino acids of MRL-DNA10 are also depicted. The amino acid change resulting from the single nucleotide difference in position 99 (GGG vs. TGG) from the BALB/c J_{H} 3 sequence is indicated by an asterisk. Complementarity-determining regions are boxed. Amino acid numbering is according to Kabat et al. (46). 5'TUT, 5' untranslated region; LP, leader peptide.
became C). Our findings suggest that the Igh-6 locus (IgM) of haplotypes a (BALB/c) and j (MRL/l) is highly conserved, which agrees with the apparent inability to serologically distinguish the corresponding proteins of the two haplotypes (27). Alternatively, although less likely, the MRL/l Igh-6 locus could be allotype a, requiring a recombinatorial event between an Igh\(^a\) C region and an Igh\(^j\) V region sublocus. The MRL/l strain has indeed been considered Igh\(^a\), based on earlier serological typing of the Igh-1 (IgG2a) locus (8). However, at that time the Igh\(^b\) haplotype was unknown, and Igh\(^b\) mice, including the prototype strain CBA/J, were classified as Igh\(^b\) (28). Furthermore, analysis of restriction fragment polymorphism using DNA probes corresponding to the \(\mu\) and \(\alpha\) switch region indicated that the MRL/l C region locus is \(j\), not \(a\).\(^3\)

**Structural Genetic Elements Creating the MRL-DNA10 L Chain.** 1–2% of the transformants in the MRL-DNA10 cDNA library hybridized to the L\(_{\alpha}\) C region probe. Restriction enzyme analysis of 10 clones suggested the presence of cDNA copies to a single L\(_{\alpha}\) mRNA species. The nucleic acid sequence (Fig. 3) corresponded to an apparently functional L\(_{\alpha}\) mRNA species that is assumed to encode the MRL-DNA10 L chain protein. This assumption was substantiated by the complete match of 18 N-terminal amino acids with those deduced from the nucleic acid sequence.

The V\(_{\alpha}\) gene segment–encoded portion (5′ untranslated region, leader peptide, amino acids 1–95) showed 96% similarity to a V\(_{\alpha}1\) L chain cDNA sequence corresponding to a hybridoma-derived anti-(Glu\(^{66}\)Ala\(^{66}\)Tyr\(^{10}\))\(_n\) (GAT) antibody from a BALB/c mouse (10) (Fig. 3). Comparisons at the protein level revealed complete identity with the TEPC-105 L chain sequence up to position 95 (not shown). Therefore, homologous V\(_{\alpha}\) genes, or at least V\(_{\alpha}\) genes from a single family (V\(_{\alpha}1\)), encoded the L chain of the MRL-DNA10 anti-DNA autoantibody, the TEPC-105 immunoglobulin, and antibodies with specificity for exogenous antigens (e.g., GAT and digoxin) (10, 11). Expression of the TEPC-105 L chain protein is controlled by the Igk-Ef2 locus, which is linked to the Igk-Ef1 and Ly-3 locus on chromosome 6. Mouse strains of the Igk-Ef2\(^a\) type (prototype BALB/c) express this L chain, whereas Igk-Ef2\(^b\) strains (prototype NZB) do not (30), suggesting that MRL/l mice are Igk-Ef2\(^a\).

The MRL-DNA10 J\(_\gamma\) segment differed from the BALB/c derived J\(_\gamma 4\) sequence by two nucleotides. The difference in codon 103 (AAA instead of AAG) was silent, whereas that in codon 100 (ACG instead of TCG) led to an amino acid substitution, i.e., threonine instead of serine (Fig. 3). These differences could result from somatic mutations or, less likely, reflect polymorphism at this locus. The C region and the complete 3′ untranslated region (data not shown) were identical to a published genomic BALB/c sequence (31) except for a single nucleotide difference in codon 205, i.e., ATC instead of ATT, a silent change. This observation further supports the notion that the C locus may be essentially monomorphic.

Figure 3. Nucleotide sequence and deduced amino acid sequence of the MRL-DNA10 L chain V region compared with that of a BALB/c (Igh*) anti-GAT hybridoma antibody L chain cDNA sequence (G8 Cal.7). Dotted lines signify identity with the MRL-DNA10 nucleotide sequence. Amino acids differing from the deduced G8 Cal.7 sequence are indicated by asterisks, as are base and amino acid differences from the BALB/c J.4 sequence (not shown). Complementarity-determining regions are boxed. Amino acid numbering is according to Kabat et al. (46). 5'UT, untranslated region; LP, leader peptide.
Discussion

The main goal of this study, which is part of an attempt to delineate the genetic origin of the major SLE-associated autoantibodies, was to identify and analyze the structural genetic elements (V, D, J, and C gene segments) encoding a monoclonal anti-DNA autoantibody (MRL-DNA10) from the lupus-prone MRL/l mouse strain. We report the molecular cloning and complete nucleotide sequence of cDNA copies corresponding to its H and L chain mRNAs. This is, to our knowledge, the first autoantibody nucleic acid sequence reported. The respective clones, which correspond to functional full-length copies of both H and L chain messages, have been synthesized in a eukaryotic expression vector system and, therefore, permit future functional studies, including study of the effect of site-specific mutagenesis and expression in transgenic mice.

Comparison of the structural elements generating the MRL-DNA10 autoantibody with published immunoglobulins revealed that the 3′ untranslated, C, and J regions of both H and L chain were nearly identical to corresponding sequences of non-autoantibody IgM molecules. The putative V H gene segment was very similar to V H genes of the NP antibody family, and the L chain most likely used a V gene of the V 1 group that is also expressed in antibodies to exogenous antigens. Only the D segment could not be classified, and could therefore represent an autoantibody-specific structural element. However, novel D segments have been found in non-autoantibody molecules and may be attributed to either high levels of somatic mutation, allotypic differences, or uncharacterized D families in the murine genome (5). Furthermore, additional monoclonal murine anti-DNA and anti-IgG autoantibodies presently analyzed in our laboratory use D segments resembling known D sequences. This MRL/l autoantibody, therefore, derived from structural genetic elements closely related to those encoding antibodies to foreign antigens in normal mice.

The extent to which the murine anti-DNA response is restricted by the genetic elements described here, and the identity and numbers of other V H and V L genes that might be involved in generation of anti-DNA autoantibodies, remains to be determined. Idiotype analysis, which may indicate the degree of murine anti-DNA autoantibody restriction, has yielded controversial results. Common reactive idiotypes have been observed on anti-DNA autoantibodies from MRL/l (32) and (NZB × W) F 1 mice (33), suggesting that a limited number of V genes may encode most anti-DNA antibodies. Conversely, Tron et al. (34) observed a large amount of idiotypic diversity amongst such antibodies from the (NZB × W) F 1 strain. Structural studies to address this question directly are, for the most part, lacking. Amino-terminal sequences of two (NZB × W) F 1-derived monoclonal anti-DNA H chains suggested that one may originate from a V H gene of the S107 gene family, which is involved in the immune response to PC, whereas the other resembles the H chain of a DNA-binding NZB myeloma protein with undetermined V gene origin (35). The latter H chain paired with an L chain whose amino-terminal sequence resembled that of an anti-PC antibody of the V 8 group. Hence, the V H and V L genes giving rise to these and the MRL-DNA10 anti-DNA autoantibody were probably derived from different families. Nucleic acid sequence analysis of several other MRL/l and (NZB × W) F 1-derived
monoclonal anti-DNA autoantibodies presently performed in our laboratory further support a multigenic origin of the murine anti-DNA response.

The genetic capacity to produce autoantibodies against a variety of self-antigens, and hence the presence of the respective structural genes in healthy individuals and experimental animals, is strongly suggested by numerous observations: autoantibodies occur in the sera of humans (reviewed in 36) and mice (37) during viral infections, and can be induced in normal background mice by bacterial products (38) or introduction of genes such as the autosomal recessive \textit{lpr} gene (39). Autoantibody-secreting hybridomas have been recovered from fusions using tonsillar lymphoid cells from a nonautoimmune donor (40) or normal mouse splenocytes (41). Furthermore, DNA from lupus and normal mice of the same \textit{Igh} haplotypes resulted in essentially identical restriction enzyme patterns with labeled DNA probes corresponding to various \textit{V} \textit{\alpha} gene families, suggesting shared Ig \textit{V} \textit{\alpha} loci.

The above experiments leave little doubt about the presence in normal individuals of the structural genes to express autoantibodies. They do not, however, clarify whether autoantibodies derive from a specific subset of \textit{V} \textit{\gamma} and/or \textit{V} \textit{\delta} germline genes or use the same structural genetic elements as antibodies to exogenous antigens. This question has relevance for possible mechanisms controlling autoantibody expression. In our study, we report that a monoclonal IgM, anti-DNA autoantibody from a murine model of systemic autoimmunity derives from structural genetic elements highly related and possibly identical to those coding for antibodies to foreign antigens. If this observation, which is further supported by the presence of autoantibody idiotypes on non-autoantibody molecules (42), can be generalized, anti-self specificity could result from certain combinations of individual elements or entire chains and/or somatic mutations of antibodies originally raised against exogenous antigens. The crossreaction of a monoclonal anti-DNA antibody with NP (this report), the N-terminal amino acid sequence similarity between human monoclonal lupus autoantibodies and anti-\textit{Klebsiella} antibodies (43), and the induction of anti-DNA autoantibodies by immunization with cardiolipin (44) favor this concept. Further direct support comes from a recent report (45) on the generation in vitro of an antibody with DNA specificity by somatic point mutation in the H chain gene of the PC-binding myeloma S107. Molecular analysis of additional monoclonal anti-DNA and anti-IgG autoantibodies, currently in progress in our laboratory, will hopefully lead to more definitive conclusions regarding the genetic origin of lupus-associated autoantibodies.

Summary

The mRNAs encoding heavy and light chains of a hybridoma-derived monoclonal IgM, anti-DNA autoantibody from lupus-prone MRL/Mp-lpr/lpr mice (Igh\textsuperscript{\textregistered}) have been transcribed into cDNA copies and molecularly cloned, and their complete nucleotide sequences have been determined. The mRNA for the heavy chain variable region, including leader peptide and 5' untranslated region, is transcribed from a heavy chain variable region (\textit{V} \textit{\alpha}) gene closely related (and possibly allelic) to \textit{V} \textit{\alpha} genes of the C57BL/6 (Igh\textsuperscript{\textregistered}) nitrophenyl antibody family. The deduced amino acid sequence corresponding to the light chain variable
region of this autoantibody shows extensive similarities with non-autoantibody molecules of the V,1 group, suggesting a common variable gene origin. The joining segments, constant regions, and 3′ untranslated regions of both the heavy and light chain mRNAs are nearly identical to corresponding sequences of non-autoantibodies from normal mice. Our findings suggest that this anti-DNA autoantibody originated from the same germline repertoire as antibodies to exogenous antigens.

We thank Drs. S. J. Francis, N. R. Klinman, P. R. Morrow, S. C. Riley, P. A. Singer, and G. J. Sutcliff for helpful discussions, Dr. B. F. Tack for providing amino acid sequencing facilities, Mr. R. S. Balderas for expert technical assistance, and Ms. M. K. Occhipinti for editing and manuscript production.

References


