Molecular Immunologic Strategies to Identify Antigens and B-Cell Responses Unique to Multiple Sclerosis

Donald H. Gilden, MD; Mark P. Burgoon, PhD; B. K. Kleinschmidt-DeMasters, MD; R. Anthony Williamson, PhD; Omar Ghausi, BS; Dennis R. Burton, PhD; Gregory P. Owens, PhD

Identification of the causative agent of multiple sclerosis (MS) has long eluded investigators and has become the “Holy Grail” of researchers in the field. The immune response in cerebrospinal fluid of patients with MS, indicated by an increased IgG level and the presence of specific oligoclonal bands after electrophoresis, strongly parallels that found in various infectious diseases of the central nervous system. To understand the nature of B-lymphocyte activation in MS, 4 laboratories studied the antigen-binding regions of antibodies found in MS brain demyelinating plaques and cerebrospinal fluid. Each analysis revealed (1) limited germline expression, results not expected for a random bystander response; (2) features consistent with a specific antigen–targeted process; and (3) the clonal expansion of populations of B lymphocytes in MS. The screening of libraries expressing protein products derived from chronic MS plaque messenger RNA with antibodies purified from plaques, cerebrospinal fluid, or serum of patients with MS has thus far not revealed the antigenic target(s) of the MS antibody response. Because putative MS antigens could be in low abundance, the screening of large libraries of random peptides expressed on phage surfaces might offer an alternative approach to identify peptide sequences recognized by MS antibodies. New sophisticated molecular immunologic techniques described herein should enhance our ability to identify putative antigen(s) targets in MS.

Normal cerebrospinal fluid (CSF) has an IgG content that is less than 13% of total protein. In multiple sclerosis (MS), there is an unexplained elevation of IgG in the CSF to 15% to 30%, visualized as oligoclonal bands (OGBs) after electrophoresis. Besides MS, OGBs in CSF are typically found in patients with infectious disorders of the central nervous system (CNS), such as neurosyphilis, tuberculous and fungal meningitis, and subacute sclerosing panencephalitis (SSPE) caused by measles virus. In these diseases, OGBs have been shown to be antibodies directed against the disease-causing organism. The specificity of OGBs in MS CSF and brain is unknown, but their presence suggests the possibility of an underlying infection. This review focuses on analyses of intrathecally synthesized oligoclonal IgG to identify specific antigens in MS brain against which the humoral immune response might be directed.

Exactly which tissue to target to identify the putative MS antigen is problematic. Frozen tissue samples of acute MS brain plaques would be the best choice for analysis, but they are usually unavailable. Even the definition of what are the earliest actively demyelinating plaques is uncertain. Attempts to identify a specific and uniform temporal stage–dependent cascade of specific inflammatory cells and immune mediators in demyelinating lesions have been problematic because immunohistochemical analyses of MS plaques have revealed great heterogeneity. The presence of macrophages expressing the MRP-14 protein (a marker of presumed early macrophage activity) seems to iden-
identify the earliest phases of myelin breakdown in some patients (John Prineas, MD, oral communication, December 6, 1999). Targeting plaque tissues for T cells is also useful because T cells appear early in lesions, followed by B-cell influx, but both may be transient and hence are less reliable markers of acute hyperacute lesions.

In the late 1990s, 4 laboratories independently used molecular immunologic and biologic strategies and techniques to study the humoral immune response in MS. Two groups studied MS brain and 2 studied MS CSF. Each laboratory used equivalent strategies that extracted RNA from brain or CSF, reverse transcribed the RNA into complementary DNA (cDNA), amplified the cDNA with primers specific for immunoglobulin heavy-chain genes, cloned the amplified IgG-specific DNA, and sequenced and analyzed the expressed genes. We describe their findings, which were remarkably similar, and their significance, with implications for further study of the humoral immune response in MS.

ANTIBODY STRUCTURE AND FUNCTION

A brief review of antibody structure and function serves as a starting point for understanding recent studies that examined the humoral immune response in MS. Antibody molecules are proteins, known collectively as immunoglobulins, and are produced by activated B cells. Immunoglobulin antibodies bind specifically to particular molecules known as antigens. Immunoglobulin binding can neutralize pathogens or mark them for uptake and destruction by phagocytes. There are 5 main classes of immunoglobulins (IgM, IgD, IgA, IgG, and IgE) that are distinguished by the immunoglobulin heavy-chain constant region. IgG is the most abundant class of antibody in human serum, human CSF, and the brain of patients with MS.

Although all antibodies have the same overall structure, each antibody molecule has unique sequence differences that confer specificity for a particular antigen. Figure 1A shows the 2 identical heavy- and 2 identical light-chain regions of IgG. Each heavy- and light-chain region contains V (variable) and C (constant) domains. One pair of heavy- and light-chain V regions at the “arms” of the IgG molecule binds to a specific antigen. Thus, each intact IgG molecule has 2 binding sites. Within the V regions of the heavy and light chains, there are 3 hypervariable regions (1, 2, and 3), the most diverse of which is the third region. The rest of the V domains between the hypervariable regions are termed the framework regions (FR1, FR2, FR3, and FR4).

The framework regions provide the structural backbone of the antigen-binding domain. When the $V_h$ and $V_l$ domains pair in the antibody molecule, the hypervariable loops from each domain are brought together to form the antigen-binding site. Because the 3 hypervariable loops constitute the binding site for antigen and determine specificity by forming a surface complementary to the antigen, these hypervariable loops are termed the complementarity-determining regions (CDRs) and are denoted CDR1, CDR2, and CDR3. The combination of heavy- and light-chain CDRs determines the final antigen specificity. Thus, one way in which the immune system is able to generate antibodies of different specificities is by generating different combinations of heavy- and light-chain variable regions.

The complete collection of antibody specificities available within an individual is known as the antibody repertoire and in humans consists of as many as $10^{11}$ or more different antibody molecules. The entire antibody repertoire is generated by recombination events involving limited numbers of heavy- and light-chain genetic elements. The heavy-chain V region ($V_h$) is more diverse and is generated by recombination of a heavy-chain variable segment, a D segment, and a J segment during B-cell development. In humans, there are approximately 51 functional $V_h$ segments, 30 functional D segments, and 6 functional J segments. Recombination of these segments provides an enormous number of antibody combinations. Whereas the $V_h$ segment determines CDR1 and CDR2, the entire D segment and the amino terminal end of the J segment (Figure 1B) encodes the heavy-chain (H) CDR3. Because the recombination mechanism that joins these segments is not precise and additional nucleotides can be added or deleted at the V-D and D-J junctions during recombination, the HCDR3 sequence is the most diverse. In fact, the chances of 2 different B cells forming identical CDR3 sequences during development is so remote that the HCDR3 sequence can be used as a clonal marker. Thus, in the $V_h$ repertoire analyses we review herein, any $V_h$ sequences that share the same or almost identical CDR3...
sequence, but contain other discrete sequence differences, are considered clonally related variants. These variants have originated from a common B-cell ancestor that was activated and underwent clonal expansion and somatic mutation (detailed in the following section).

Any substance can elicit an antibody response, and the response even to a simple antigen is diverse, comprising many different antibody molecules, each with a unique affinity and fine specificity. The diversity of the antibody repertoire reflects not only the fact that there are separately inherited genes for each different antibody chain, but, more important, that somatic mutation occurs. Somatic mutation refers to the alterations that occur in the V region sequences of activated B cells after antigenic stimulation.

EVIDENCE FOR AN ANTIGEN-DRIVEN IMMUNE RESPONSE IN MS

The target of the humoral response in MS is unknown, but several possibilities exist. First, B-cell activation may result from antigenic stimulation (foreign or self) targeted to a specific molecule(s). Second, activation could result from a nonconventional mechanism such as B-cell superantigen stimulation. B-cell superantigens are proteins, usually viral or bacterial, that activate B cells polyclonally irrespective of antigen specificity but dependent on some feature of variable domain framework architecture. Third, activation could be a random bystander effect of the inflammatory response seen in MS plaques. The first studies to distinguish between random and targeted humoral immune responses were based on the premise that OGBs are synthesized within the brain because plasma cells and large quantities of IgG messenger RNA (mRNA) are present in MS plaques but not in normal human brain white matter. A strategy using the polymerase chain reaction yielded a representative sampling of the IgG VH repertoire expressed in MS plaques that was sequenced and analyzed (Figure 1). The VH repertoire shared many hallmarks associated with antigen-driven responses, including a skewed distribution of VH germline segments in multiple brain plaques compared with peripheral blood, extensive somatic mutation, preferential accumulation of amino acid replacement mutations in the antigen-binding CDRs, and the presence of distinct sequence differences in some overexpressed populations (clonal variants), indicative of clonal expansion.

A second, larger study compared the VH repertoire in MS plaques with that found in SSPE in which CSF and brain OGBs result from a continuous antigen-driven response directed against measles virus. First, when the IgG sequences in SSPE brain were expressed as antibody, they were shown to be directed against measles virus antigen. Second, when the VH sequences in MS and SSPE brain were further analyzed, sequence variation was found. Instead of the 98%-100% homology normally found among VH genes in different individuals (human polymorphisms account for 1%-2% variation), the average homology with germlines was only 92%, indicative of extensive somatic mutation in MS and SSPE brains. Third, specific VH sequences in SSPE and MS brains were overrepresented, and many sequences preferentially accumulated replacement mutations in CDRs relative to the framework regions. Comparison of VH family and germline usage showed that patients with SSPE or MS had a response characteristic of a targeted, antigen-driven process. Similar conclusions regarding the humoral response in MS were also reached in studies that analyzed the length and amino acid structure in the CDR3 regions of IgG in 10 additional MS brains. Compared with the peripheral VH repertoire in normal blood, most MS brains displayed skewed VH gene usage, arguing against a random B-cell response. Furthermore, 2 independent studies of the VH repertoire in CSF of different patients with MS also revealed the same features. Overall, these molecular immunologic findings in MS, where the antigen is unknown, are similar to the humoral response in SSPE, where the antibodies are known to be directed against measles virus. If the antibody responses in MS were due to superantigen stimulation, consistent restricted use of the same germline segments would be expected. While VH4 germline segments predominated in the first MS brain analyzed, analyses of additional MS brains revealed a unique response profile in each brain, again inconsistent with superantigen stimulation but consistent with an antigen-driven, targeted response.

cDNA EXPRESSION LIBRARY SCREENING

The IgG antibody in MS brain and CSF may be directed against an antigen (protein) crucially involved in the disease process. Furthermore, this putative MS antigen has probably been translated from mRNA expressed in MS brain. Thus, it would be valuable to analyze all the mRNAs expressed in MS brain to survey their encoded proteins for antibody specificity. To do so, all the RNA extracted from MS brain is first reverse transcribed into cDNA. The cDNA is then cloned into a suitable cloning vector to create a cDNA library. This approach uses special cloning vectors, called λ expression vectors, in which the cloned DNA is transcribed into the complementary mRNA, which in turn is translated into the encoded protein.

When replicate nitrocellulose filters are prepared from a cDNA library constructed in a λ expression vector, fusion proteins expressed from each individual clone are bound to the nitrocellulose filter. Thousands of individual clones can be screened on 1 filter. The replicate filter can be screened by procedures capable of detecting specific fusion proteins. For example, an antibody specific for a protein of interest, eg, IgG from MS brain or CSF, can be incubated with replicate filters of a λ cDNA expression library. If one of the λ clones expresses a fusion protein that includes the region of the protein bound by the antibody, antibody molecules will bind to the filter at the position of that specific clone. After washing the filter to remove unbound antibody, the position of the
specific clone is detected by incubation with a second radioactively labeled antibody that recognizes the first antibody, followed by autoradiography of the filter.

We screened cDNA expression libraries prepared from acute and chronic plaques in MS brain. Two cDNA libraries prepared from pathologically verified chronic plaques were screened with pooled MS CSF containing OGBs and with IgG extracted from MS CSF or MS serum, but they did not reveal MS-specific antigens. Furthermore, immunoblotting with pooled MS CSF IgG or IgG eluted from MS plaque-periplaque white matter did not reveal myelin proteins or any MS-specific antigens in MS or normal brain white matter.

Another group searched for possible CNS-specific autoantigens in MS by screening cDNA expression libraries generated from an oligodendrocyte precursor cell line. A library screened with pooled CSF from 54 patients with MS detected 6 positive clones, of which 5 contained a common 7-amino acid sequence highly homologous to the translation product of an alu repeat sequence. Subsequent screening with serum and CSF samples from patients with MS showed that nearly half reacted with these so-called alu peptides. Although alu sequences compose 5% of the human genome, the expression of these encoded peptides in the brain or their relevance to the pathogenesis of MS is not known.

Identification of the antibodies in the CNS and CSF of patients with MS that correlate with disease is a logical first step in discerning their corresponding antigens. Recombinant antibody technology combined with phage display offers a novel strategy to identify disease-relevant IgG and their corresponding antigens in MS. Phage display technology involves the fusion of foreign DNA fragments to genes encoding filamentous phage coat proteins. A phage (abbreviated from bacteriophage) is a virus that infects bacteria. Antibody-phage libraries rapidly generate specific high-affinity antibodies from immune donors. Recombinant antibody libraries have been selected against purified or crude antigenic preparations, yielding antibodies that specifically recognize a variety of viral pathogens and autoantigens. In antibody libraries, sequences encoding IgG Fab (antibody binding) fragments from target tissue (eg, MS brain) are amplified by polymerase chain reaction and cloned sequentially into a phage display vector, yielding random combinations of IgG heavy- and light-chain binding fragments. Each Fab fragment is fused to the gene encoding phage coat protein III, which generally affords monovalent display of the recombinant antibody on the phage surface. This approach has been used to create phage libraries containing diverse repertoires of more than 10^7 antibody clones. Repeated testing of the library against a chosen antigenic preparation (“panning”) leads to the recovery of monoclonal antibodies that react specifically with the selecting antigen. Successive panning rounds also reduce the diversity of phage-Fabs binding to the antigen to yield antibody clones with the highest affinity for antigen. Figure 2 outlines the production and application of phage antibody libraries in chronic CNS infections. This technology can be applied to MS by cloning the antibody response localized in and around acute MS plaques and in CSF of patients with MS. Antigenic prepa-
tions of MS brain tissues can then be used to select specific disease-associated antibodies within the phage library.

The feasibility of this approach was demonstrated in a test system using an antibody Fab-phage library prepared from gray and white matter of SSPE brain. Panning against lysates of measles virus-infected cells\(^{14}\) or sections of SSPE brain\(^{15}\) selected Fabs that immunostained SSPE brain tissue and reacted with measles virus–specific antigens. Thus, IgG mRNA expressed in brain during a chronic CNS infection can be used to generate high-affinity antibodies that recognize antigens of the disease-causing pathogen.\(^{15}\)

Applying recombinant antibody technology combined with panning on single antibodies. So far, one potential pitfall of this technique to characterize the humoral immune response in human disease. Antibodies from the brain or CSF of patients with MS can be used to identify their corresponding protein epitopes, however rare, from a peptide library. Epitope libraries contain exhaustive arrays of short nucleotide sequences that encode random amino acids at each of 6 or more positions. These sequences are displayed on the surfaces of phage.\(^ {17}\) Libraries containing all possible hexapeptides, ie, the functional size limit of most antibody epitopes, would contain 1 billion combinations and challenge the upper size limit for a workable library. But libraries of several hundred million independent clones with longer peptide insertions of about 15 amino acids also contain enough potential hexapeptide sequences and are easier to construct.

Using panning protocols similar to the Fab display libraries described previously, approximately \(10^{12}\) phage-displayed peptides can be selected over immobilized antibody preparations. The peptide regions in bound phage are readily sequenced to reveal consensus peptides that correspond to the antibody’s target. Comparison of consensus peptide epitopes to databases can then reveal clues to the in vivo antigen against which the antibody was generated, particularly as databases expand. The peptide library technique is most useful to identify linear epitopes of antibodies, and even perhaps competing epitopes of antibodies against nonpeptide targets. One potential pitfall of this technique is the detection of mimotopes, peptides that mimic the original antigen binding but do not share sequence homology with the native antigen and therefore are not useful in database searches. Because this technique can be founded by polyclonal antibody responses, it is best suited for panning on single antibodies. So far, one group has reported\(^ {18}\) that antibody from MS CSF identified a 5–amino acid consensus sequence in 5 of 14 patients. The same peptide sequence is present in the Epstein-Barr virus nuclear antigen and in \(\alpha\)-crystallin, a heat-shock protein active in some MS myelin extracts. The relevance of these findings to MS is still being investigated.

**FUTURE DIRECTIONS**

In addition to the strategies and techniques described herein that have been applied to MS, another immunologic strategy for pathogen detection and identification is based on the use of new gene microchip array techniques. Screening which genes are activated in response to a specific infectious agent may result in characteristic host molecular signatures generated in response to infection or toxins. Specific microbial stimuli may elicit a characteristic gene expression response profile that can serve as a diagnostic signature of infection. The collection and analysis of expression response profiles from cells exposed to infectious agents and toxins in vitro, and expression response profiles from peripheral blood mononuclear cells of healthy individuals, uses custom-designed human cDNA microarrays with 18000 elements, representing more than 15000 expressed genes. Early results suggest that expression profiles can discriminate between different members of the same bacterial genus and between specific virulence factors, eg, toxins.\(^ {19}\)

Multiple studies have elucidated the nature of the humoral immune response in MS and have revealed the hallmarks of an antigen-driven response. Now, the presence of a novel antigen can be investigated with the most sophisticated and sensitive techniques of molecular immunology.

*Accepted for publication June 29, 2000.*

*This work was supported in part by grants NS 32623 (Dr Gilden) and AI 39162 (Dr Burton) from the Public Health Service, Bethesda, Md.*

*We thank Marina Hoffman for editorial review and Cathy Allen for preparing the manuscript.*

Corresponding author and reprints: Donald H. Gilden, MD, Department of Neurology, University of Colorado Health Sciences Center, 4200 E Ninth Ave, Mail Stop B182, Denver, CO 80262 (e-mail: don.gilden@uchsc.edu).
REFERENCES


Correction
