MEDICAL GRAND ROUNDS

THE GENETIC ORIGIN OF AUTOANTIBODIES

J. DONALD CAPRA, M.D.

Professor of Microbiology and Internal Medicine

June 30, 1988 11:00 a.m.

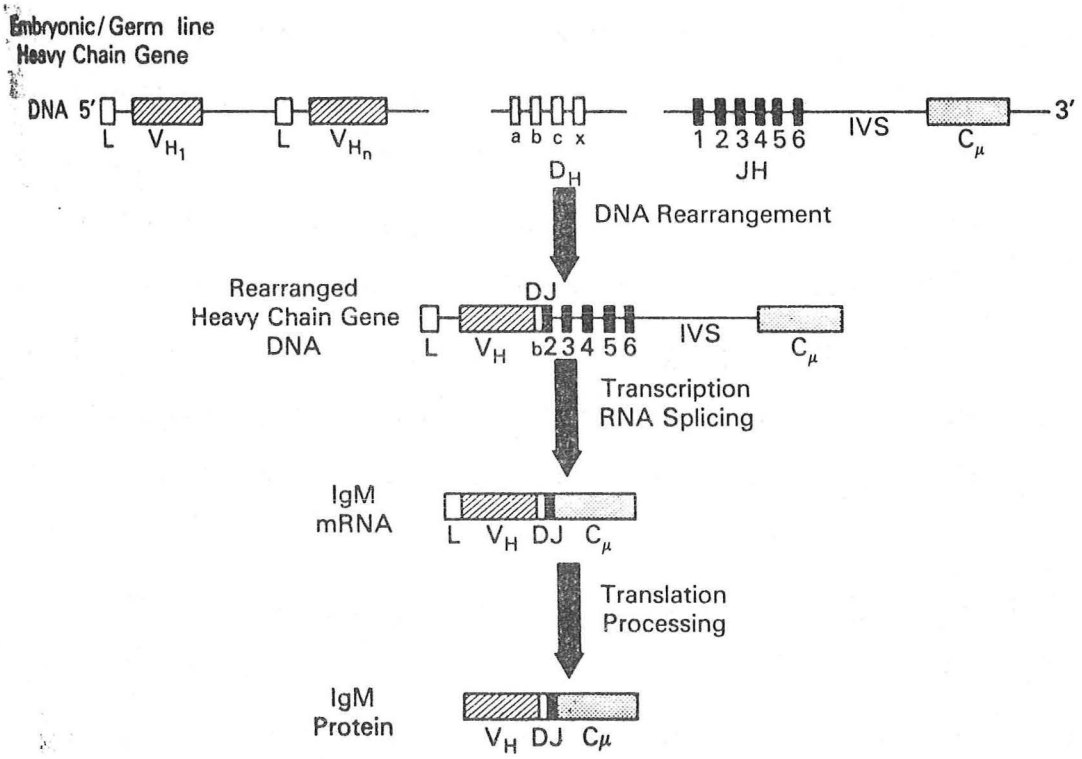
I. INTRODUCTION

These Grand Rounds will concern three basic issues that remain in the field of human antibody diversity. First, we will deal with issues concerning the genetic organization of the human immunoglobulin variable region complex. Second, we will address the issue of polymorphism within this gene complex, particularly as it might relate to human disease - fundamentally, we wish to know if the difference between individuals who acquire immunologic disease and those who do not might be encoded in the genes encoding the variable regions of immunoglobulin molecules. And, finally, we will address the question posed by the title of these Grand Rounds. "What is the genetic origin of human autoantibodies?" Two basic theories are extant. The first teaches that autoantibodies can form in anyone and are unrelated to the genetics of the immunoglobulin variable region complex, that autoantibodies may or may not be germline encoded and that they represent bystanders in these diseases. The second hypothesis teaches that autoantibodies are direct copies of germline genes and that only individuals carrying certain germline genes are capable of making certain autoantibodies and, thereby, acquiring certain autoimmune diseases. While progress has been made in understanding the first two questions that I have posed, I should warn you at the outset that the last dilemma remains although it is my firm prediction that an answer will be available within the very near future (Sanz and Capra, 1988).

II. THE ORGANIZATION OF THE HUMAN VARIABLE REGION GENES - THE V_H REGION.

While studies on the myeloma proteins in man provided the earliest insights into the structure of antibody molecules (Kunkel, 1965;

Edelman, and Gally, 1964), there has been a paucity of information available on the genes encoding human antibodies. Studies in the murine system have far exceeded the human and information that is available, for example in the inbred Balb/c mouse, outdistances the molecular analyses of the human V_H locus by a factor of 100 to 1. The reasons for this are complicated but largely revolve around the ease at producing murine hybridomas and the difficulty at immortalizing human B cells. Thus, in the murine system, it has been relatively easy to produce monoclonal antibodies of nearly any specificity, immortalize cell lines and through recombinant DNA techniques, isolate the genes for a whole series of antibodies, utilize these as molecular probes into the genome, and thereby map the entire murine V_H region. Let me digress for a moment and review that data with you.



The immunoglobulin variable region gene complex encoding the heavy chain consists of approximately 1,000 germline V gene segments, a minimum of ten D segments and 4-6 J segments (Rathbun, et al., 1988a). These are thought to recombine in a relatively random way during ontogeny so that the familiar Chinese menu approach to antibody diversity is obvious and that one can choose one from column A, one from column B and one from column C and make 1,000 X 10 X 5 or approximately 50,000 different variable regions from only 1,015 germline gene segments. As each of these segments joins through recombination, additional nucleotide variability is introduced such that with every joint there can be a minimum of three additional amino acids increasing this diversity by a factor of 10. Thus, the germline, while encoding only 1,015 gene segments, can give rise to nearly a half million different structures. Somatic mutation works on top of this to provide an unlimited repertoire of variable region structures which can deal with literally any antigenic challenge.

A HALF MILLION ANTIBODY V_H REGIONS FROM 1000 GENE SEGMENTS

1000 V_H Genes X 10 D_H Genes X 5 J_H Genes = 50,000 Combinations

At Each Joint ~3 Different Structures are Possible , thus 50,000 x 3 x 3 = 450,000 V_H Regions

The organization of the murine V_H locus is unusual. As the genes were sequenced, they fell into ten families. Families are operationally defined as being approximately 80% homologous at the nucleotide level and out-of-family members are less than 70% homologous. Few genes fall in between. An alternative definition of a gene family is that under standard conditions of Southern filter hybridization, genes from different families don't cross hybridize but "within" family genes do. The families are remarkably different in size varying from two to over 500 and antibodies of particular specificities are often found in only a single gene family. Thus, many of antibodies to carbohydrate antigens are members of the X24 gene family while many antibodies to small organic haptens are in the J558 gene family (Rathbun, et al., 1988b).

The organization of these families is also unusual. By and large the evidence is that the families are clustered. Thus, two of the four members of the S107 family are physically linked. These clusters of related families make a certain amount of sense in that they allow for the alignment of the DNA during cell division and would presumably act to prevent or to discourage non-homologous recombination. It has been assumed that the human V_H locus is organized in a similar fashion.

THE MURINE V_H MAP

5'-3609-J558-(VGAM3-8, J606, S107)-3660-

(X24, Q52, 7183)-3'-D_H

A series of most intriguing observations were made by several investigators almost simultaneously concerning the utilization of the murine V_H genes during ontogeny. The bulk of the experimental evidence suggests that the genes that are most proximal to D and J are rearranged early (so-called "proximal utilization model"). Thus, if one analyzes the repertoire of antibodies in the fetus one sees the utilization of the Q52 and 7183 gene families while neonates utilize the S107 and J606 gene families and only in adults does one see the utilization of J558 and 3609, the more distal V_H gene families.

There is also a fair amount of experimental evidence to suggest that the vast majority of autoantibodies derived from these "D-J proximal" gene families. Thus, for example, a series of rheumatoid factors isolated from murine strains with autoimmune disease are largely members of the V_H7183 family (Manheimer-Lory, et al., 1986) and a group of anti-thyroglobulin antibodies are members of 7183 and Q52 gene families (Zanetti, et al., 1983). It was anticipated that the human autoimmune system would parallel those of the mouse as there have been so many similarities between, for example, human lupus erythematosus and the several murine models.

I should not leave this murine model without emphasizing that there are some disquieting conundrums that continue to be bothersome: Not all of the V_H families are clustered (Rathbun, et al., 1988a), not all autoantibodies come from the J proximal families (Wu and Page, 1986), not all of the studies on ontogeny agree with the proximal utilization model.

THE HUMAN V_H LOCUS

The human heavy chain complex is located on the long arm of the 14th chromosome. There have been relatively recent developments that have led

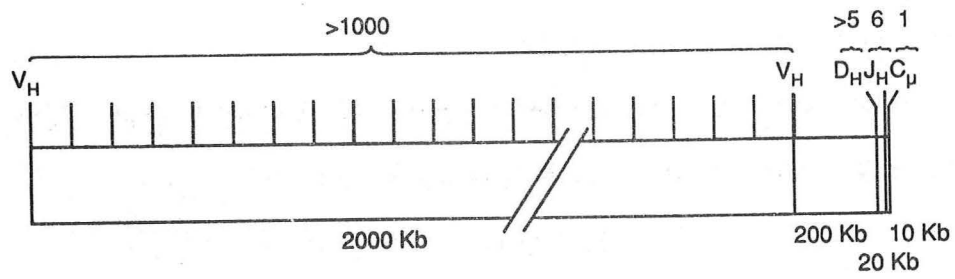
to an explosion of information concerning the human V_H complex. First, slowly but surely, human hybridoma technology is taking hold. Second, EBV transformation of human peripheral blood lymphocytes has allowed for the immortalization of human B cells for some time. However, only recently has it been appreciated that while such cell lines did not provide in general enough protein for amino acid sequencing, they provide ample mRNA for cDNA cloning, and third, the emerging interest in human B cell leukemias and lymphomas as sources of rearranged immunoglobulin V_H genes. All three of these sources have led to the isolation of a series of human V_H cDNA probes which have then been used to isolate genomic clones.

Coupled to this analysis has been the commitment of two major laboratories - those of Fred Alt at Columbia University and Tasuku Honjo in Japan to isolate, sequence and link a large number of human V_H genes (Lee, et al., 1987; Alt, et al., 1987; Berman, et al., 1988). These two laboratories have cloned and sequenced approximately 50 human V_H germline gene segments and perhaps the rest of the scientific community has contributed a dozen more (Shen, et al., 1987; Humphries, et al., 1988). Thus, we have had available within the last six months, our first glimpses into the organization and structure of the human V_H locus. Recall that the first 100 myeloma protein sequences came from humans and they were easily divided into three V_H subgroups or families. Essentially all the proteins could be easily classified as such and it was widely assumed that unlike the mouse which consisted of ten V_H families, the human contained but three. Gradually, however, as the work of Alt, Honjo and Tucker at the genomic level and Perlmutter (Schroeder, et al., 1987) and our own laboratory (Sanz, et al., 1988) at the expressed antibody level, was analyzed, it was appreciated that the human V_H complex consisted of a

minimum of six and perhaps eight different V_H families. These newly discovered V_H families will be a significant focus of these Grounds Rounds as it is through these newly discovered, relatively small V_H families that a great deal of insight into the organization and utilization of the human V_H locus has been deduced.

Gene probes for V_HI, II and III came relatively early, largely from cloning the rearranged genes from B cell lymphomas (Ben-Neria, et al., 1981; Rechavi, et al., 1983). The V_HIV family was discovered accidentally. Honjo was linking human V_H gene segments in a cosmid walk and uncovered a sequence that was remarkably different than any previously described human V_H sequence. At the time, no expressed antibodies similar to this had been detected. A molecular probe of this new V_H gene (subsequently called V_HIV) was used in Southern filter hybridization experiments and a new series of hybridization fragments was detected thereby acknowledging this new V_H gene family. A remarkable feature of the family was its relatively small size. V_HI, II and III had consisted of over 50 restriction fragments. V_HIV on the other hand, consisted of 5-10 genes. Only recently, in the Perlmutter laboratory, has the cDNA from a fetal liver library been isolated that confirms that V_HIV genes can indeed be expressed and in our own laboratory we have isolated a series of autoantibodies that utilize the V_HIV gene segment. We will return to this issue later.

HUMAN Igh COMPLEX



The $V_H V$ family was discovered here at Southwestern by Philip Tucker (Shen, et al., 1987; Humphries, et al., 1988). Cloning the genes from a family in which several members had chronic lymphatic leukemia a gene was isolated and sequenced that had only modest relationship to any previously described V_H family. When used as a hybridization probe, it revealed a new set of restriction fragments that had not been detected with $V_H I$, II, III, and IV probes. In addition, like $V_H IV$, $V_H V$ was quite small comprising only three restriction fragments. These three genes were cloned and sequenced. Two turned out to be pseudogenes and, therefore, the $V_H V$ gene family consists of a single functional gene! Subsequently, in our own laboratory, two expressed antibodies utilizing $V_H V$ have been isolated, one, an anti-insulin antibody from a patient with diabetes who has been chronically treated with insulin (Sanz, et al., submitted). More importantly, as we will discuss later, Dr. Tucker and his colleagues have demonstrated that this gene is rearranged preferentially in approximately

30% of patients with certain lymphocytic leukemias (Humphries, et al., 1988).

The V_HVI family was discovered independently by Alt and Perlmutter utilizing different approaches - the former by genomic cloning, the latter by cDNA cloning. This family is also small consisting of only 1-3 members. To date, no antibody of defined specificity has been found which expresses a V_HVI gene segment.

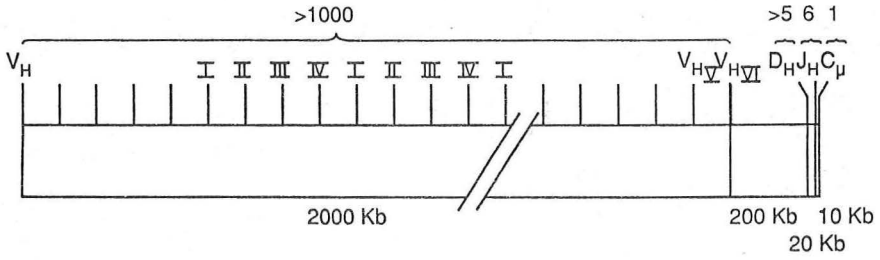
There are additional single genes that are suggestive of additional families but we will stop at this point.

To recapitulate then, within the last six months, there have been two major advances. First, the discovery of three additional gene families (V_HIV, V, VI) and, second, these new families appear to be small and preferentially utilized in both B cell tumors and autoantibodies.

The second important observation that has come largely from the studies of Alt and Honjo is that unlike the mouse where the V_H families are organized into discreet family clusters, the human V_H families are interspersed. Honjo, for example, has reported single cosmid clusters containing members of four different human V_H families and Tucker and Alt have independently reported phage clones concerning members of two different V_H families. Phage clones with two V_H genes of differing families have rarely been described in the mouse. These data suggest that the human V_H locus might be organized quite differently from the murine V_H complex. Parenthetically, studies from Kay Knight's laboratory in Chicago (Galarda, et al., 1985) suggest that the rabbit V_H region is organized much like the human with vastly different V_H structures adjacent in the genome. The implications for this distinct orientation are obvious for any extension of the murine "proximal utilization models" to

the human if the V_H genes are organized in an interspersed fashion. If proximal utilization is the rule, then the human fetus and neonate would have the opportunity to call upon vastly different V_H genes to deal with both internal and external antigens rather than, like the mouse, call on rather similar genes for these purposes.

HUMAN Igh COMPLEX



Studies by pulse field electrophoresis and deletion mapping have suggested that the newly discovered V_H families are J proximal. The V_HV gene which is commonly rearranged in acute and chronic lymphocytic leukemia has been linked by pulse field electrophoresis to D and J_H (Humphries, et al., 1988). This demonstrated physical linkage of the variable region to the constant region of human immunoglobulins for the first time. These data suggested that of all the human V_H genes, V_HV was closest to D and J_H. This provided a rational explanation for the common rearrangement of V_HV in B cell tumors. Later, Alt documented that the single V_HVI gene was even closer to D and probably was separated by only 25-70 kb (Berman, et al., 1988). Subsequent experiments suggested that the order of the genes is from D_H:V_HVI, then V_HV, then V_HIV - at least some V_HIV genes. The fact that the human V_H genes are so

interspersed makes further analysis of this kind extremely complicated as virtually every hybridization probe gives multiple bands even with pulse field electrophoresis.

SOURCES OF MATERIAL FOR HUMAN ANTIBODY STUDIES

- Human hybridomas
- EBV transformed B lymphocytes
- B cell leukemias and lymphomas

Studies done in our own laboratory utilizing those EBV transformed cell lines in which but a single immunoglobulin V_H complex is present (the other having been lost during the immortalization of the cell line) confirms this general organization.

Thus, in summary, the human V_H locus is organized quite differently than the mouse. Rather than the genes appearing in clusters that are related by family, they seem to be interspersed. However, because of these newly discovered V_H families are small, while not all of the members are J proximal several of them are and both V_HIV, V_HV and V_HVI, the newly discovered V_H families are among the closest to the D and J segments. Many molecular biologists feel that proximity is a significant determinant in recombination and, as such, these V_H gene segments that are close to D and J may have a more important role in human physiology than the size of these families would suggest.

III. THE HUMAN V_K LOCUS

The human kappa genes are located on chromosome 2 and like the heavy chain complex consists of a number of variable and J gene segments. It differs from the heavy chain in that the light chain locus contains no D gene segments and, therefore, the diversity mechanisms are somewhat more limited; that is, V segments and J segments can form several different structures and junctional diversity between them can amplify this perhaps by a factor of three or four. Since it is assumed that there are approximately 100 V kappa gene segments and five J kappa gene segments, 100 X 5 = 500 X 5 = 2500 different V kappa variable regions can be generated. Unlike the situation with the heavy chain locus where there are multiple constant region genes (mu, delta, gamma, alpha, epsilon) and the same variable region can "switch" to different constant region genes depending upon the needs of the organism, in the kappa locus there is but a single constant region gene. Thus, in many respects, the kappa locus is considerably simpler than the heavy chain locus.

Most of what we know about human kappa genes and proteins comes from the work of two prominent German scientists. The first human kappa chain was sequenced by Hilschmann 23 years ago (Hilschmann and Craig, 1965) and represented the first immunoglobulin chain sequenced. This was from a patient with multiple myeloma with Bence-Jones proteinuria. It was these first two sequences done in the laboratory of Lyman Craig at the Rockefeller University that ushered in the new era of structural analyses of immunoglobulin molecules. From this work came insights into the variable and constant regions of immunoglobulins as well as the first hints of polymorphism. Over the next twenty years Hilschmann and his laboratory sequenced over 100 human kappa chains from myeloma proteins and through it

defined the four families of human kappa chains, defined the variations within these chains and provided the framework for the vast majority of molecular biological studies that followed.

Within the last five years, another prominent German scientist, Hans Zachau, has committed his laboratory to the complete cloning and sequencing of the entire human V kappa locus. To date, over 100 V kappa genes have been cloned and sequenced, and most of these have been mapped in overlapping cosmid clusters (Pohlenz, et al., 1987). From these analyses have come the following general conclusions. First, the $V_K I$, $V_K II$, $V_K III$, and $V_K IV$ gene segments are intersperced. Thus, there is none of the clustering that is seen in the murine V_H families. Second, the orientation of these genes vis-a-vis the J and C regions varies with some genes oriented 3' to 5' and others 5' to 3'. This argues that the mechanism of V-J joining can either be by deletion or inversion. Third, there are a minimum of two and perhaps three segments of the V kappa locus that have undergone inversion and duplication such that the familial relatedness of a series of genes in one orientation is nearly identical to the sequence and organization in the reverse orientation. This organization, which has important implications for the evolution of the kappa locus, is also of considerable importance in the expression of kappa genes. Finally, those genes that are most "J proximal" have been mapped although studies on their utilization are only beginning. The 100 genes that Zachau's laboratory have mapped exist on a cluster of approximately 1000 kb of DNA.

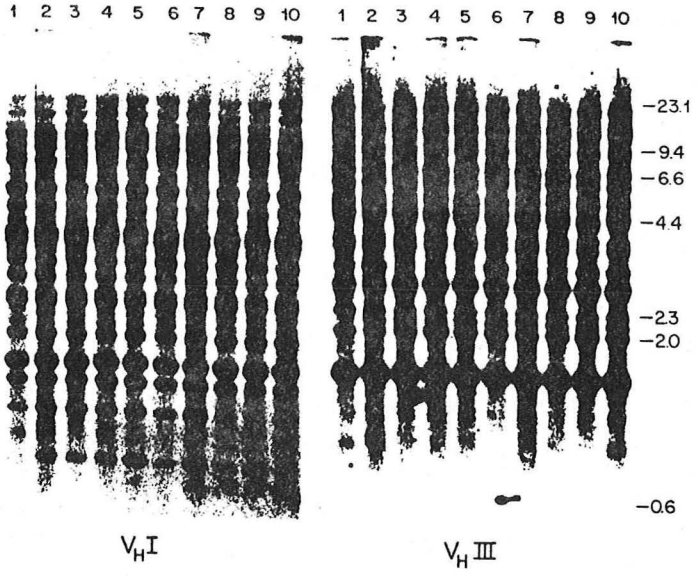
POLYMORPHISM OF HUMAN V_H GENE SEGMENTS

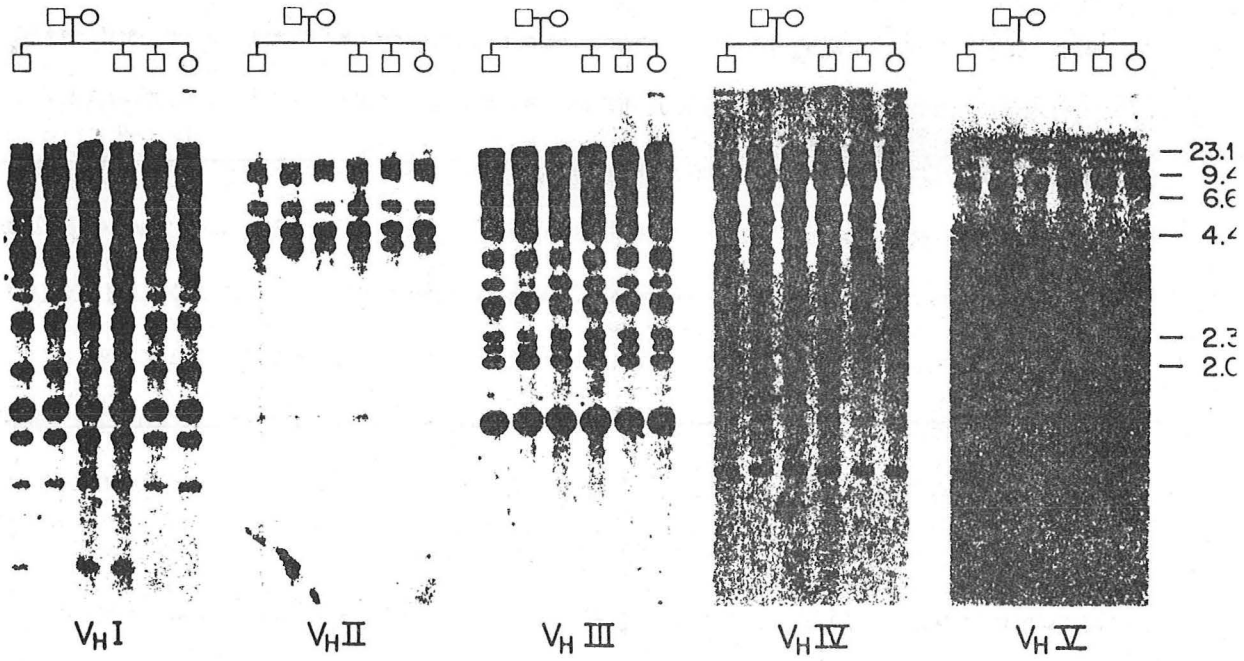
The genetic factors on human immunoglobulins were described over 30 years ago by some of the pioneers in the field of immunology including Grubb, Oudin, and Kunkel. Genetic markers are well known in the constant regions of human immunoglobulin kappa chains as well as the heavy chains of all immunoglobulin classes and subclasses. The nomenclature of these markers is that the chain of designation is listed first such as kappa, gamma, alpha; designated K, G, or A; followed by the small letter m to indicate "marker" and then an arabic letter. Thus, kappa light chains contain markers known as Km1 and 2, and similarly the markers on the gamma chain are called Gm1, 2, etc. all the way to approximately 35. The genetic markers were instrumental in the infancy of human immunology in understanding the diallelic nature of the expression of these molecules, phenomena such as allelic exclusion and the inheritance of the whole cluster. With few exceptions, most genetic markers in the human system are single amino acid interchanges that occur on exposed sites of the molecule rendering them antigenic. Most antisera that are used to distinguish these markers have historically come from patients who are multiply transfused or from multiparous women. More recently, mouse monoclonal antibodies are available that distinguish all of these genetic markers along with the classes and subclasses so that the analysis of immunoglobulins over the last decade has become a routine laboratory test.

Genetic markers in the variable region have been extremely difficult to define. Serologically, while there are a number of reports describing these, few have been reproduced and none has reached the point of general acceptance. The protein chemists have not been particularly helpful in this regard as the amounts of material required were such that few multiple

structures from the same individual were ever defined. It has fallen to the molecular biologist to define such polymorphisms and within the last few years using genetic probes, such polymorphisms have been described both in the kappa and heavy chain locus. At the present time, the vast majority of these polymorphisms are defined by so-called RFLPs or restriction fragment length polymorphisms, a system that has been described at These Grand Rounds repeatedly over the last year. The majority of these RFLPs are defined by polymorphic sites in introns outside of the coding regions and, therefore, their physiologic significance remains in question.

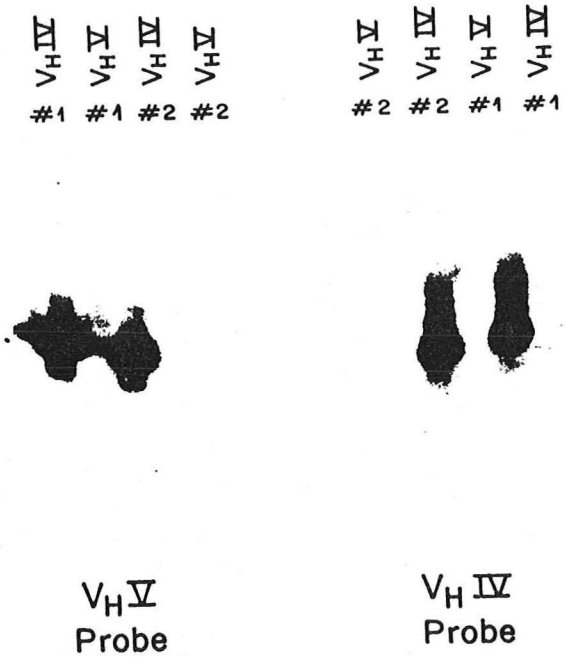
The difficulty in addressing the issue of polymorphism in the human population at the immunoglobulin variable region locus is complicated by the multiplicity of germline genes at both the V_K and V_H loci. It becomes extremely difficult in looking at the DNA of two unrelated individuals to find a single restriction fragment that is truly allelic and since large family studies are extremely difficult in the human population and inbreeding is the exception rather than the rule, little progress has been made in this arena.





The discovery of the V_HIV, V and VI families of immunoglobulin variable region genes that I described earlier provides an important opportunity to address this issue. Since, for example, there is but a single V_HV gene that is functional in the human genome, it has been possible to isolate this gene from unrelated individuals and determine its nucleotide sequence. Dr. Tucker's and my own laboratory have done this at both the genomic and the expressed level and have detected such polymorphisms. These studies represent the first clear examples of germline encoded polymorphism in the variable regions of immunoglobulin genes. They provide the context for the next subject of These Grand Rounds

which concerns the relationships of polymorphisms to human disease.

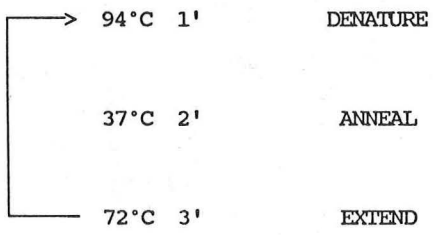


Let me take a few moments to explain how this study was done. First it was appreciated that there was but a single functional V_HV gene. This gene had been cloned and sequenced in the Tucker laboratory and we had the availability of not only that sequence but the powerful new technology of polymerase chain reaction which has been described here within the last six months by both Mike Brown and Graham Smith. In this procedure, oligonucleotides are constructed such that only the V_HV genes can be amplified. Amplification is done in such a way that restriction sites are placed within the oligonucleotide used for amplification. After amplification, appropriate restriction endonuclease digestion leads to the generation of an easy cloning system and these genes can be placed into plasmid and phagemid vectors for direct sequencing. Several individuals in our laboratory volunteered their DNA for this study and the results are shown in the figure below.

POLYMERASE CHAIN REACTION

1μM Primers

1ng DNA



What is striking is that while each of these sequences are remarkably similar, They are indeed different. These differences are both coding and non-coding such that some are silent and some lead to amino acid interchanges. These data strongly argue that the single functional V_HV gene of different individuals encodes different variable region structures. What we don't know at this time is whether this variation leads to functionally different molecules. Every indication is that this is the case. These and other data like them provide a rational explanation for differences in immune response upon antigenic challenge of different humans who may be identical at their major histocompatibility complex genes.

Let us digress for a moment. From studies that were done by Benaceraff over two decades ago and which led to the awarding of the Nobel Prize in Medicine about five years ago, it is known that the immune response in guinea pigs and mice is highly dependent upon the genes

in the major histocompatibility complex. These genes were originally called IR (immune response) genes and we now know them as the class II antigens for the DR, DQ and DP genes of the human histocompatibility complex. However, a large body of experimental evidence suggests that other genes also control immune responses. In mice it is known that variations in immune response can map to the variable regions of both the heavy and light polypeptide chains. This has been an extremely difficult issue to address in man because of the obvious experimental constraints. However, there are a number of situations where siblings that are HLA identical exhibit different immune responses particularly to the well known vaccines that are administered such as pneumococcal, hepatitis, and DPT.

POLYMORPHIC IMMUNOLOGICALLY RELEVANT MOLECULES AND THEIR CHROMOSOMAL LOCATIONS

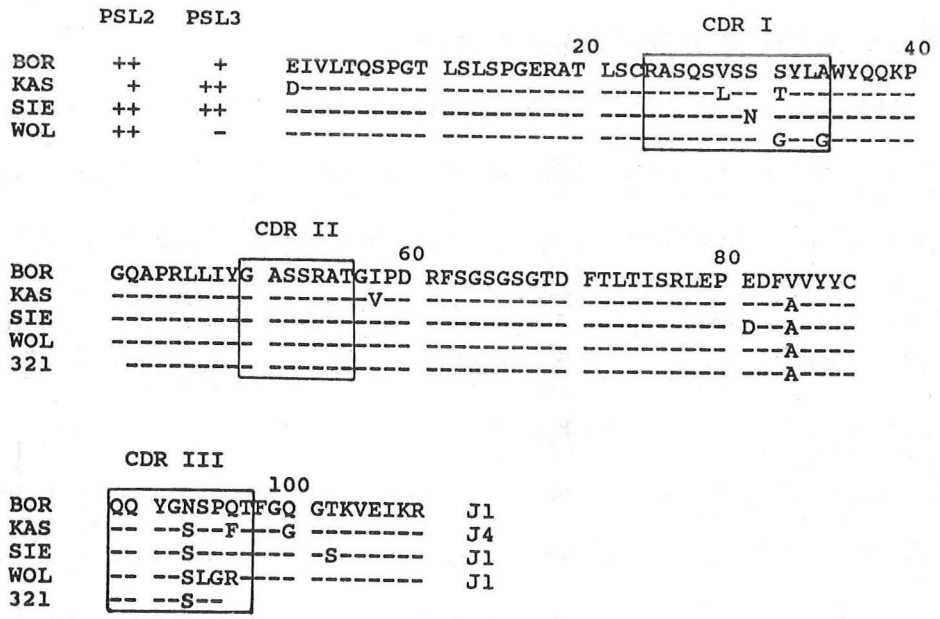
Molecules	Chains	Location
Immunoglobulin	Heavy	14
	Kappa	2
	Lambda	22
T Cell Receptor	Alpha-Beta	14,7
	Gamma-Delta	7
Histocompatibility	Class I	6
	Class II	6
	Beta-2M	15

It has been impossible because of the small numbers of patients involved to accurately map the location of these variations in immune response. These data provide a rational explanation for these findings. If two individuals have different V_H genes, for example, particularly genes that have differences in their coding regions as has been illustrated here, one could anticipate a differential immune reponse; that is, one individual could respond with a higher affinity antibody than another. These provide

important insights into how autoimmune disease may be distributed within a family and gives insight into how a "second" gene outside of the MHC could be implicated in human disease. Recall that at my last Medical Grand Rounds, I proposed that variation at the T cell receptor locus could impact on human disease. There are now a wide variety of experimental systems and a host of publications concerning the relationship between polymorphisms of the human T cell receptor and celiac disease, rheumatoid arthritis, autoimmune thyroiditis, insulin dependent diabetes mellitus and myesthenia gravis. In my view, studies of the human V_H and V kappa locus will provide similar genetic predispositions to human disease.

THE GENETIC ORIGIN OF HUMAN AUTOANTIBODIES

Between 1968 and 1974, several papers were published describing cross-idiotypic specificities among human autoantibodies, specifically, cold agglutinins (Williams, et al., 1968) and monoclonal rheumatoid factors (Kunkel, et al., 1973; 1974). These reports provided some of the first insights into the genetic origin of autoantibodies demonstrating that all members of one of the cross-idiotypic groups of human rheumatoid factors contained V_KIIIb light chains by serologic analysis. These discoveries were followed by primary structural analyses of human rheumatoid factors which defined their structural relatedness (Capra and Kehoe, 1974; Andrews and Capra, 1981). It is important to recall that these early studies (as well as most studies until recently), employed monoclonal rheumatoid factors derived from patients with B cell malignancies. Few of these patients had manifestations of rheumatoid disease. The relationship of these monoclonal rheumatoid factors to the polyclonal rheumatoid factors characteristic of rheumatoid arthritis has never been entirely established. Nonetheless, these studies were interpreted to suggest that proteins with similar idiotypes contained similar structures in their hypervariable regions. Not appreciated at the time (because relatively few human proteins had been sequenced) was that the extraordinary similarity of the structures of these molecules suggested that they might derive from similar if not identical germline genes.



During the past ten years, a large body of experimental evidence has accumulated, largely emanating from the laboratory of Dennis Carson at Scripps Clinic defining in precise serologic and structural terms the relationships and genetic origins of this most interesting group of human antibodies. The most recent of these studies was just published in the Journal of Immunology (Crowley, et al., 1988). It reports the interrelatedness of a large panel of monoclonal paraproteins with a series of precisely defined serologic reagents largely directed toward monoclonal human rheumatoid factors. The study concludes that the vast majority of rheumatoid factors bear the serologic markers of V_KIIIb light chains although certain important exceptions exist.

Our present understanding of the genetic origin of human autoantibodies represents a significant advance for modern immunochemistry. Carson and his group have employed monoclonal antibodies to define the serologic determinants on this group of human

autoantibodies. In addition, they have produced anti-peptide sera to the various hypervariable regions of both the heavy and light polypeptide chains of the various serologic groupings of these molecules. The reactivity of these anti-peptide sera with polyclonal rheumatoid factors, monoclonal rheumatoid factors, with large panels of paraproteins and in normal sera demonstrates the precision at which one can define antibody molecules at the present time. These studies have nicely dovetailed with the cloning of the germline light chain gene segments that give rise to both V_KIIIa and V_KIIIb rheumatoid factors; thus contributing greatly to our understanding of the molecular basis of autoantibodies and autoimmunity (Kipps, et al., 1977; Chen, et al., 1987a, 1987b; Silverman, et al., 1988; Goldfien, et al., 1987).

The genetic origin of these molecules has long remained an enigma. This question lies at the heart of a fundamental understanding of autoimmunity. Do autoantibodies directly derive from germline genes? Do all of us carry the same complement of immunoglobulin germline genes? Is the distinction between individuals who get autoimmune disease and those who do not based on differences in their immunoglobulin V_H or V_L genes or are somatic and/or regulatory factors involved? These issues are being addressed in a fundamental way by a number of scientists in several different systems.

Studies of murine rheumatoid factors by Weigert and his group (Schlomchik, et al., 1987a) as well as by the Scripps group of Dixon and Theofilopoulous (Aguado, et al., 1987) as well as on murine anti-DNA antibodies by Barrett's group (Trepicchio, et al., 1987) provide evidence that there are not fundamental structural differences between autoantibodies that arise in various murine models of autoimmunity and

autoantibodies that can be induced and/or selected in normal mice. More importantly, the sequences of many such autoantibodies show numerous somatic mutations, the distribution of which suggests positive selection by antigen (Schlomchik, et al., 1987b).

In the human system it has been much more difficult to address these issues largely because of the difficulty in producing human-human hybridomas and except for the model system of mixed cryoglobulinemia, there are relatively few human situations where monoclonal autoantibodies are easily available. However, within the last year, several groups have developed systems to address this issue. Livneh, et al. (1987), for example defined the so-called 8.12 idiotype among human lupus antibodies () and recently, Hoch and Schwaber (1987) identified and sequenced the V_H gene elements encoding a human anti-DNA antibody. Dersimonian, et al., (1987) have published an important paper concerning the structure of various human anti-DNA hybridomas documenting that the structures of two human anti-DNA antibody V_H regions, one derived from a patient with leprosy and one derived from a patient with lupus were absolutely identical. These data argue that autoantibodies need not require the somatic mutation of a germline gene. An important difference between the majority of human versus mouse studies is that, by and large, the murine antibodies sequenced were of the IgG class while the human antibodies were IgM.

Recent studies from our own laboratory confirm these observations. In collaboration with Paolo Casali and Abner Notkins, we have examined the V_H nucleotide structures of several monoclonal polyreactive antibodies and have found some that had identical nucleotide sequences although derived from genetically distinct individuals. Similarly, in collaborative

studies with Howard Dang and Norman Talal, we have determined the complete V_H structure of an anti-Sm antibody derived from a patient with systemic lupus erythematosus (Sanz, et al., submitted). In this instance, the nucleotide sequence of the V_H gene segment was identical to a cDNA clone recently published from Perlmutter's group and obtained from a fetal liver cDNA library (Schroeder, et al., 1987). This result shows: 1) that the V_H gene used by this SLE specific autoantibody is most likely germline encoded; 2) that these genes are without significant polymorphism in the general population and, therefore, could play an important physiological role since evolutionary pressure is acting undoubtedly to preserve these structures, and; 3) that at least some of these autoreactive V_H genes are expressed early in the development of the B cell repertoire. Collectively these studies further suggest that autoantibodies derived from normal individuals (those without disease) and autoantibodies derived from patients with disease may be structurally identical.

That the human V_H and V_L genes used in many autoantibodies derive from relatively small V_H and V_L gene families may be of crutial importance. Several of the polyreactive antibodies are the first examples of functional V_HIV genes and an anti-insulin antibody studied in our laboratory by Tom Thomas is the first example of a functional V_HV gene. Kipps, et al. (1987) have shown that a significant number of patients with chronic lymphatic leukemia utilize the V_KIIIb light chain gene, and Humphries, et al. (1988), reported that almost a third of patients with lymphocytic leukemia rearrange genes of the V_HV variable region family. This latter observation is particularly remarkable in view of the fact that there are a maximum of three V_HV germline genes representing at most 3% of the entire human V_H gene repertoire. The predominant utilization of some V_H families in specific B cell tumors along with the predominance of the "classical" families (V_HI, V_HII, V_HIII) in paraproteins produced by plasma cell tumors arising in the bone marrow suggests that B cells from distinct subsets at different stages of maturation and/or in different compartments with different microenvironments may utilize different sets of V_H genes.

Out of these studies have come two fundamental points. First, the polymorphism in the human appears to be dramatically less than the polymorphism that distinguishes the various inbred murine strains and, second, at least among the IgM autoantibodies, their V_H and V_L gene segments appear to be direct copies of the germline genes.

Why, then, do autoantibodies in patients with autoimmune disease appear "aggressive" while "normal" autoantibodies appear relatively benign? Two main structural explanations come to mind: (1) The D segments of these two groups of antibodies appear significantly different and have no known germline counterparts. This suggests that either there are additional germline D segments or novel mechanisms are involved in the generation of D segments in autoantibodies (Meek, et al. 1987). (2) The pathogenic autoantibodies still remain largely unknown and few have been subjected to structural analysis. It may well be that somatic events operate upon physiological autoantibodies and transform them either by increasing their affinity or by changing the idiotype, which in turn may allow them to escape idiotype control.

Should the genetic origin of autoantibodies be no different in normals and in patients with autoimmune disease, we are left with the relatively uninspiring conclusion that autoimmunity has no genetic component amongst the immunoglobulin genes. This would fly in the face of a large body of data suggesting that idiotypes of both rheumatoid factors and anti-DNA antibodies are heritable as well as considerable statistical evidence suggesting a second or a third genetic system in addition to the MHC is involved in human autoimmunity. One way to reconcile the information is to postulate that the difference lies in complex regulatory pathways of the immune system (timing of expression in ontogeny, selection

of T cell repertoire, etc.). All those factors, along with MHC products and environmental agents could interact to expand in autoimmune patients, clones that in normals are also present but down regulated.

REFERENCES

Aguado, M.T., R.S. Balderas, R.L. Rubin, M.A. Duchosal, R. Kofler, B.K. Birshtein, D.S. Secher, F.J. Dixon and A.N. Theofilopoulos. 1987. J. Immunol. 139: 1080.

Alt, F.W., T.K. Blackwell, and G.D. Yancopoulos. 1987. Development of the primary antibody repertoire. Science 238: 1079.
28.

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.

Ben-Neria, Y., J.B. Cohen, G. Rechavi, R. Zakut, and D. Givol. 1981. Polymorphism of germline immunoglobulin V_H genes correlates with allotype and idiotype markers. Eur. J. Immunol. 11: 1017.

Berman, J.E., S.J. Mellis, R. Pollock, C. Smith, H.Y. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C. Cantor and F.W. Alt. 1988. Content and organization of the human Ig V_H locus: Definition of three new V_H families and linkage to the Ig C_H locus. EMBO J. 7: 727.

Capra, J.D., and J.M. Kehoe. 1974. Structure of antibodies with shared idiotypy: the complete sequence of the heavy chain variable regions of two IgM antigamma globulins. Proc. Natl. Acad. Sci. USA 71: 4032.

Capra, J.D., and D.G. Klapper. 1976. Complete amino acid sequence of the variable domains of two human IgM anti-gamma globulins (Lay/Pom) with shared idiotypic specificities. Scand. J. Immunol. 5: 677.

Chen, P.P., D.L. Robbins, F.R. Jirik, T.J. Kipps, and D.A. Carson. 1987a. Isolation and characterization of a light chain variable region gene for human rheumatoid factors. J. Exp. Med. 166: 1900.

Chen, P.P., K. Albrandt, T.J. Kipps, V. Radoux, F.-T. Liu, and D.A. Carson. 1987b. Isolation and characterization of human V_KIII germline genes: Implications for the molecular basis of human V_KIII light chain diversity. J. Immunol. 139: 1727.

Crowley, J.J., R.D. Goldfien, P.A. Carson, and S. Fong 1988. Incidence of three cross reactive idiotypes on human rheumatoid factor paraproteins. J. Immunol. 140: 3411.

Dersimonian, H., R.S. Schwartz, K.J. Barrett, and B.D. Stollar. 1987. Relationship of human variable region heavy chain germline genes to genes encoding anti-DNA autoantibodies. J. Immunol. 139: 2496.

Diamond, B., and M.D. Scharff. 1984. Somatic mutation of the T15 heavy chain gene gives rise to an antibody with autoantibody specificity. Proc. Natl. Acad. Sci. USA 81: 5841.

Edelman, G.M., and J.A. Gally. 1964. The model for the 7S antibody molecule. Proc. Natl. Acad. Sci. USA 51: 846.

Gallarda, J.L., K.S., Gleason, and K.L. Knight. 1985. J. Immunol. 135: 4222.

Goldfien, R.D., P. Chen, T.J. Kipps, G.Starkebaum, J.G. Heitzmann, V. Radoux, S. Fong and D.A. Carson. 1987. Genetic analysis of human B cell hybridomas expressing a cross-reactive idiotype. J. Immunol. 138: 940.

Hilschmann, N., and L.C. Craig. 1965. Amino acid sequence studies with Bence-Jones proteins. Proc. Natl. Acad. Sci. USA 53: 1403.

Hoch, S., and J. Schwaber. 1987. Identification and sequence of the V_H gene elements encoding a human anti-DNA antibody. J. Immunol. 130: 1689.

Humphries, C.G., A. Shen, W.A. Kuziel, J.D. Capra, F.R. Blattner, and P.W. Tucker. 1988. Characterization of a new human immunoglobulin V_H family that shows preferential rearrangement in immature B cell tumors. Nature, in press.

Kipps, T.J., S. Fong, E. Tomhave, P.P. Chen, R.D. Goldfien, and D.A. Carson. 1987. High-frequency expression of a conserved K light-chain variable-region gene in chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA 84: 2916.

Kunkel., H.G. 1965. Myeloma Proteins and Antibodies. In: The Harvey Lectures 59, Academic Press, p. 219.

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-gamma-globulin activity. J. Exp. Med. 137: 331.

Kunkel, H.G., R.J. Winchester, F.G. Joslin, and J.D. Capra. 1974. Similarities in the light chains of anti-gamma-globulins showing cross-idiotypic specificities. J. Exp. Med. 139: 128.

Lee, K.H., F. Matsuda, T. Kinashi, M. Kodaira and T. Honjo. 1987. A novel family of varaible region genes of the human immunoglobulin heavy chain. J. Mol. Biol. 195: 1.

Livneh, A., J.-L. Preud'homme, A. Solomon, and B. Diamond. 1987. Preferential expression of the systemic lupus erythematosus-associated idiotype 8.12 in sera containing monoclonal immunoglobulins. J. Immunol. 139: 3730.

Manheimer-Lory, A.J., M. Monestier, B. Bellon, F.W. Alt, and C.A. Bona. 1986. Fine specificity, idiotypy and nature of cloned heavy-chain variable region genes of murine monoclonal rheumatoid factors. Proc. Natl. Acad. Sci. USA 83: 8293.

Meek, K., D. Jeske, S. Alkan, J. Urbain, and J.D. Capra. 1987. Structural characterization of syngeneic and allogeneic anti-idiotypic antibodies in the anti-arsonate system. From: Idiotypes and Disease. Eds. M. Zanetti, C.A. Bona, and F. Celada. Karger AG, Basel, page 109.

Pohlenz, H.-D., B. Straubinger, R. Thiebe, M. Pech, F.-J. Zimmer, and H.G. Zachau. 1987. The human V_K locus: Characterization of extended immunoglobulin gene regions by cosmid cloning. J. Molec. Biol. 193: 241.

Rathbun, G., J.D. Capra, and P.W. Tucker. 1988a. Organization of the murine immunoglobulin V_K complex in the inbred strains. The EMBO J. 6: 2931.

Rathbun, G., I. Sanz, K. Meek, P. Tucker, and J.D. Capra. 1988b. The molecular genetics of the arsonate idiotype system in A/J mice. Adv. in Immunol. 42: 95.

Rechavi, G., D. Ram, L. Glazer, R. Zakut, and D. Givol. 1983. Evolutionary aspects of immunoglobulin heavy chain variable region (V_H) gene subgroups. Proc. Natl. Acad. Sci. USA 80: 855-859.

Sanz, I., and J.D. Capra. 1988. Polymorphisms of immunologically relevant loci in human disease: Autoimmunity and human heavy chain variable regions. N.Y. Acad. Sci. In press.

Schroeder, H.W., J.L. Hillson, and R.M. Perlmutter. 1987. Early restriction of the human antibody repertoire. Science 238: 791.

Shen, A., C. Humphries, P. Tucker, and F. Blattner. 1987. Human heavy-chain variable region gene family nonrandomly rearranged in familial

chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA 84: 1987.

Shlomchik, M.J., A. Marshak-Rothstein, C.B. Wolfowicz, T.L. Rothstein, and M.G. Weigert. 1987a. The role of clonal selection and somatic mutation in autoimmunity. Nature 328: 805.

Shlomchik, M.J., A.H. Aucoin, D.S. Pisetsky, and M.G. Weigert. 1987b. Structure and function of anti-DNA autoantibodies derived form a single autoimmune mouse. Proc. Natl. Acad. Sci. USA 84: 9150.

Silverman, G.J., R.D. Goldfien, P. Chen, R.A. Mageed, R. Jefferis, F. Goni, B. Frangione, S. Fong, and D.A. Carson. 1987. Idiotypic and subgroup analysis of human monoclonal rheumatoid factors: Implications for structural and genetic basis of autoantibodies in humans. J. Clin. Invest., in press.

Trepicchio, W., A. Maruya, and K.J. Barrett. 1987. The heavy chain genes of a lupus anti-DNA autoantibody are encoded in the germline of a nonautoimmune strain of mouse and conserved in strains of mice polymorphic for this gene locus. J. Immunol. 139: 3139.

Williams, R., H.G. Kunkel, and J.D. Capra. Antigenic specificities related to the cold agglutinin activity of IgM globulins. Science 161: 379.

Wu, G.E., and C.J. Paige. 1986. The EMBO J. 5: 3475.

Zanetti, M., M. DeBaets, and J. Rogers. 1983. High degree of idiotype cross-reactivity among murine monoclonal antibodies to thyroglobulin. J. Immunol. 131: 2452.